AWARD NUMBER: W81XWH-11-1-0814

TITLE: Development of Technologies for Early Detection and Stratification of Breast Cancer

PRINCIPAL INVESTIGATOR: David R. Walt, PhD

CONTRACTING ORGANIZATION: Tufts University
Boston, MA 02111

REPORT DATE: October 2015

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
October 2015	Annual	09/29/2014 – 09/28/2015

4. TITLE AND SUBTITLE

Development of Technologies for Early Detection and Stratification of Breast Cancer

5a. CONTRACT NUMBER

W81XWH-11-1-0814

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Dr. David Walt, david.walt@tufts.edu; Dr. Daniel Chiu, chiu@uw.edu;
Dr. Charlotte Kuperwasser, charlotte.kuperwasser@tufts.edu;
Dr. Gail Sonenshein, gail.sonenshein@tufts.edu;
Dr. Rachel Buchsbaum, rbuchsbaum@tuftsmedicalcenter.org
email: david.walt@tufts.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Tufts University
62 Talbot Ave.
Medford, MA 02155

Tufts University School of Medicine
136 Harrison Ave.
Boston, MA 02111

University of Washington
4333 Brooklyn Ave NE
Seattle, WA 98195

Tufts Medical Center
800 Washington Street
Boston, MA 02111

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

The overall goal of this work is to develop ultra-sensitive detection techniques to identify a panel of new biomarkers and indicators with diagnostic and predictive value in breast cancer. To date, we have successfully created 13 Single Molecular Array (SiMoA) assays which will be further utilized for determining biomarker concentrations in breast cancer patients. We are screening multiple biomarkers in breast cancer patient and healthy control samples and using additional markers, such as mtDNA, to improve the sensitivity and specificity of these assays. We used PSA to demonstrate that biomarkers can be measured at ultrasensitive levels within serum using SiMoA prior to tumor formation in a mouse model. We have successfully achieved full quantification of protein molecules within single cancer cells and discovered that highly cultured cells exhibited different protein expression levels than low-passage cells. Future work will also focus on integrating ensemble Decision Aliquot Ranking (eDAR) with SiMoA to increase the throughput of the overall system.

15. SUBJECT TERMS					
Single molecule detection, cancer biomarkers, ultra-sensitive protein assays, single cells, human-in-mouse model, miRNA, circulating tumor cells					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON USAMRMC
a. REPORT	b. ABSTRACT	c. THIS PAGE	UU	41	19b. TELEPHONE NUMBER *(include area code)*
Unclassified	Unclassified	Unclassified			

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

Table of Contents

Page

Statement of Work

The overall goal of this work is to develop ultra-sensitive detection techniques in order to identify a panel of new biomarkers and indicators with diagnostic and predictive value in breast cancer. During years 1 and 2 we will develop ultra-sensitive detection techniques and apply them to identifying prospective biomarkers using the Human-in-Mouse (HIM) model of breast cancer. During years 3-5 we will extend the findings in the HIM model to validate prospective biomarkers in human subjects with breast cancer. This work is broken down into specific tasks by investigator as follows:

INTRODUCTION

Despite the recent advances made in breast cancer diagnostics and treatment, in 2013 the United States is estimated to diagnose approximately 230,000 new cases of invasive breast cancer, resulting in approximately 40,000 deaths.[1] Mammography is a powerful imaging technique for tumor detection; however, it lacks the ability to decipher benign from cancerous tumors, is unable to detect tumors smaller than 1 mm,[2] misses approximately 20% of breast cancers potentially present at the time of screening, and has an 8-10% false positive rate.[3] These drawbacks lead to inaccurate patient diagnosis, which can allow potentially fatal disease progression, or in the cases of over-treatment, unnecessary physical and emotional trauma.[4] ELISA, the most common immunoassay for measuring proteins from breast tumors, excised samples, and serum, has a lower detection limit of ~1-10 pM,[5] which is not sensitive enough to measure low abundance proteins, RNAs, and other biomarkers that could aid in the early and reliable diagnosis of cancer. There is a strikingly clear need to develop techniques capable of detecting biomarkers specific for breast cancer that will enable earlier diagnosis of disease, prediction of patient outcome, and improve therapeutic efficacy in a non-invasive manner. Our goals are to utilize ultrasensitive single molecule techniques developed in our laboratory to discover new biomarkers that meet these requirements within serum so that a simple blood test can be implemented. We are also working to characterize breast cancer biopsy samples with single cell resolution to discover the nature of the underlying heterogeneity in complex cell populations with the goal of correlating disease outcome with genotypes and phenotypes of individual cells.

BODY

David R. Walt, PhD, *Tufts University, Department of Chemistry, 62 Talbot Ave, Medford, MA 02155*

Task 1. Develop single molecule diagnostics using HIM (Human-in-Mouse) model (Years 1-2).

1a. Select approximately ten candidate markers (mt DNA, proteins, stem cell markers, etc.) in conjunction with collaborators (Months 1-6).

This task is essentially completed. See Task 1b for the list of developed assays for protein and nucleic acid biomarkers. Additional candidate markers will be added to the list as they are discovered.

1b. Develop single molecule assays for the candidate markers selected in 1a. Assays will be developed on microspheres and tested first on standard samples. Assay performance characteristics will be determined to ascertain that they will address concentration ranges of interest and requisite precision. For general protocols see: Nature Biotechnology, 2010, 28, 595-599 (Months 3-15)

Overall progress on protein biomarker assay development is summarized in **Table 1**. Table 1 compares the assay sensitivity of commercially available ELISA kits and SiMoA assays for each marker.

Table 1. Overall progress summary of breast cancer biomarker assay development.

Biomarker	ELISA LOD (pg/mL)	SiMoA LOD (pg/mL)	Fold Improvement
PR	12.5	0.3	42x
ERα	15.6	0.3	50x
LCN2	40	0.42	100x
TBX3	116	0.24	480x
CRYAB	150	30.1	5x
CYR61	3.8	0.02	190x
CA 15-3	1 U/mL	0.01 U/mL	100x
CDKN2D	10	0.10	100x
BORIS	78	10	8x
ADAM12	125	3	40x
SLUG	300	8.4	36x
SNAIL	156	5.9	26x
HER2	14.8	0.3	50x
HIF1α	50	0.015	1000x
NEDD9	1000	16.1	63x
CA 19-9	5 U/mL	0.005 U/mL	1000X
FOXC2	156	13.8	11x
VEGF	31.2	0.03	1000x

We are currently working on combining these assays into a multiplexed format for simultaneous detection of multiple markers in our diagnostic panel.

Multiplexing assays

Multiplex assays allow for in-depth analyses of blood samples without using larger sample volumes, facilitating a more high-throughput approach with minimal sample material. Dye-encoded magnetic capture beads are used to differentiate up to 10 different bead types, representing 10 different targets. Several breast cancer protein biomarkers have been selected and tested for multiplex assay development.

HIF1α, CYR61, and CA19-9 were tested for multiplexing capability. Results showed that our sensitivity for detecting HIF1α was retained in a duplexed assay with CYR61 or with CA19-9. However, CYR61 and CA19-9 were not successfully duplexed. The CA19-9 assay interfered with CYR61, as evidenced by the lack of signal in the CYR61 assay when combined with CA19-9. Furthermore, the signal generated by CA19-9 was an order of magnitude greater when paired with the CYR61 assay reagents. **Figure 1** shows the successfully duplexed assays, HIF1α/CYR61 and HIF1α/CA19-9. Further optimization will be performed on HIF1α/CYR61 in order to make the assays more compatible without compromising the LOD of either assay.

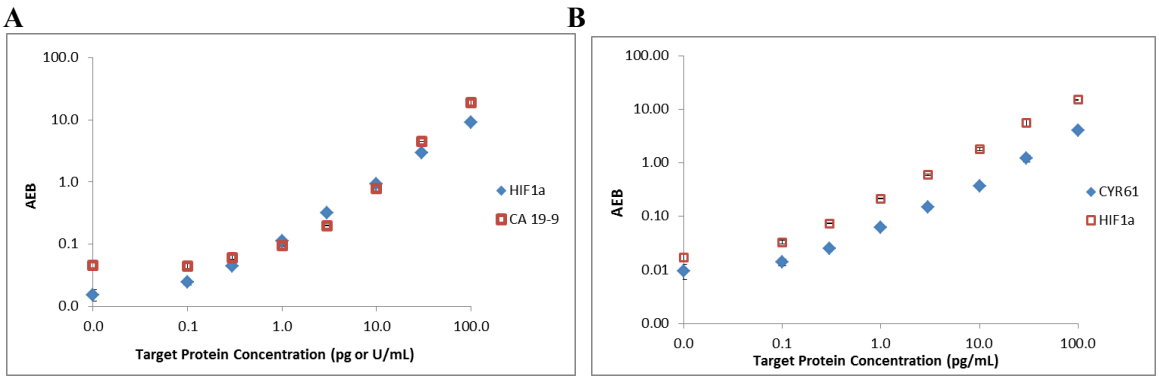


Figure 1. (A) HIF1α/CA19-9 duplex assay with LODs of 0.11 pg/mL and 0.46 U/mL, respectively. (B) HIF1α/CYR61 duplex assay with LODs of 0.039 pg/mL and 0.18 pg/mL, respectively.

HER2 and NEDD9 were tested for duplex assay development. Our sensitivity for detecting NEDD9 and HER2 were retained in a duplex assay however, all the HER2 assay signals were raised when compared to HER2 single-plex assays. From dropout experiments, the results showed that the source of the increased signals might be high concentration of the enzyme or the quality of HER2 antibodies. To avoid the HER2 signal increase, NEDD9 assays can be multiplexed with other established assays that require higher enzyme concentration. **Figure 2** shows the successfully duplexed assay with HER2 and NEDD9.

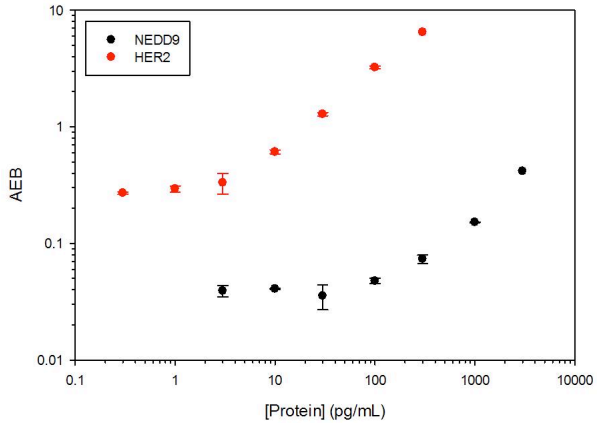


Figure 2. Duplexed detection of HER2 and NEDD9. The LOD for HER2 was 0.59 pg/mL, while the NEDD9 plex had an LOD of 37 pg/mL.

We also attempted to develop a duplex assay using HER2 and CYR61. However, the HER2 detection antibody significantly interferes with the CYR61 assay, making this combination unsuitable for multiplexing. This interference could be eliminated if the HER2 detection antibody was replaced with a HER2 SOMAmer from SOMAlogic, which is a promising substitute. The HER2 assay was successfully duplexed with PR, as shown in **Figure 3**.

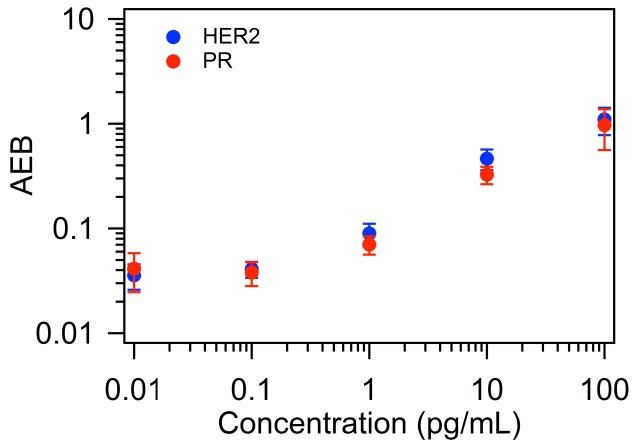


Figure 3. Duplexed detection of HER2 and PR. The LOD for HER2 was 0.3 pg/mL, while the PR plex had an LOD of 0.6 pg/mL.

Detection of microRNA as a biomarker for early stage breast cancer

Additionally, we are developing SiMoA assays for candidate nucleic acid biomarkers, specifically microRNAs (miRNAs), for diagnosing early stages of breast cancer. As a key regulator of gene expression, miRNAs are involved in virtually every biological pathway, and thus represent a very rich source of biological information. Detection of miRNAs is traditionally challenging to achieve via conventional detection methods. For example, the short length of miRNA makes it incompatible with standard PCR primers. We will achieve miRNA detection using the Single Molecule Array (SiMoA) platform. This detection method is capable of counting single molecules, has been demonstrated on a number of proteins and nucleic acids, but has not yet been successfully applied for miRNA detection. We applied two protocols to capture microRNA on paramagnetic beads: a sandwich protocol and a reverse transcription protocol (**Figure 4**). The sandwich protocol proved to be the more successful approach to detect miRNA in a highly sensitive and specific manner.

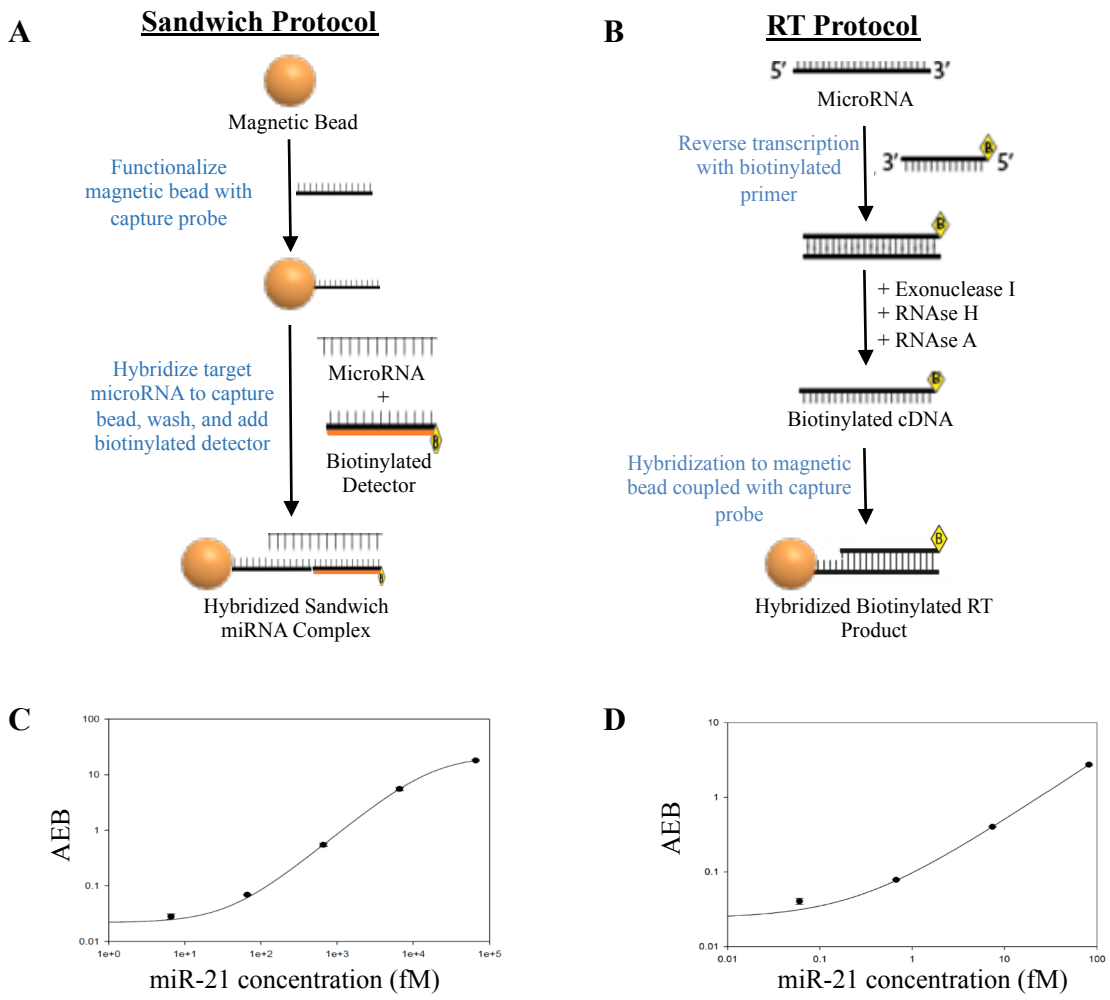


Figure 4. A) Schematic of sandwich protocol for microRNA detection. B) Schematic of reverse transcription protocol for microRNA detection. Following hybridization steps for both approaches, a streptavidin-tagged enzyme conjugate is added with fluorogenic enzyme substrate to image with SiMoA (not pictured). C) Detection of miR-21 using Sandwich Protocol, which yields a LOD of 0.66 fM. D) Detection of miR-21 using RT protocol, which yields a LOD of 30.4 fM.

Detection of homologous microRNA with single nucleotide difference

We applied a sandwich-two temperature protocol to detect homologous microRNAs that are one or two nucleotides different from each other. We chose human hsa-let-7a, hsa-let-7b and hsa-let-7c. There is only one nucleotide difference between Let-7a and Let-7c and two nucleotides difference between Let-7a and Let-7b. The nucleotide sequences are listed in **Table 2**, with the differences highlighted in red. The capture probe and detector probe are specific for Let-7b in this assay. We were successfully able to detect Let-7b at a concentration of 250 fM with only 2-3% cross-reactivity (**Figure 5**).

Table 2: Comparison between oligonucleotide sequences and melting temperatures of let 7a, Let-7b, and Let-7c.

MicroRNA	Nucleotide sequence	# of mismatched nucleotides (with respect to Let 7b)	T_m of microRNA sequences
hsa-let-7a-5p	UGAGGUAGUAGGUUGUAUAGUU	2	42° C
hsa-let-7b-5p	UGAGGUAGUAGGUUGUGUGGUU	0	49° C
hsa-let-7c-5p	UGAGGUAGUAGGUUGUAUGGUU	1	48° C

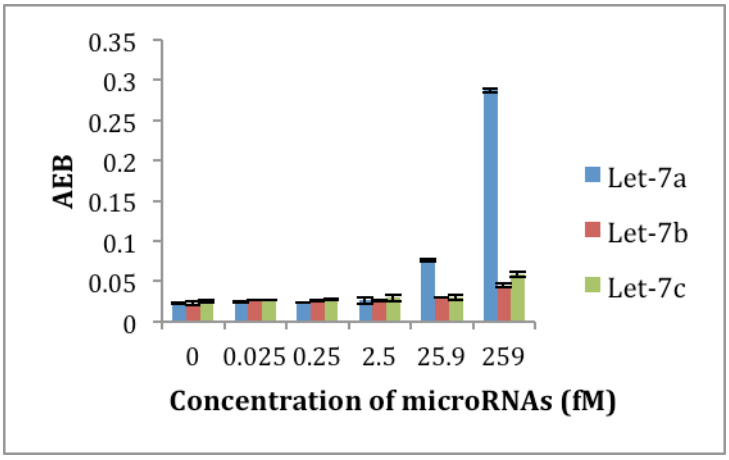


Figure 5: Cross reactivity test between homologous microRNAs – 7a, 7b, and 7c. The cross reactivity between Let-7a and Let-7b was 2%, and between Let-7a and Let-7c was 3%.

Detection of mutations in mitochondrial DNA

The Sonenshein lab has identified multiple mutations in mitochondrial DNA (mtDNA), but PCR does not have the sensitivity required to detect some of these mutations in small volumes of murine serum (provided by the Kuperwasser lab). The SiMoA platform will be applied to detect low abundance mutations of the target mtDNA sequences. A hybridization and ligation-based approach will be employed as follows: (1) the target mtDNA sequence binds to complementary DNA capture and biotinylated detection sequences, forming a sandwich complex. If the mutation of interest is present, the hybridization is a perfect match. (2) On-bead ligation is performed; only the perfectly matched sequences are ligated. (3) Target mtDNA is washed away, leaving a labeled DNA sequence for subsequent detection via enzymatic readout. This process is illustrated in **Figure 6**. Target, capture, and detection sequences are shown in **Table 3**.

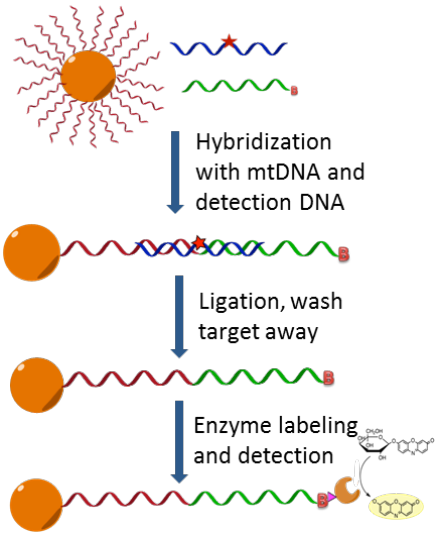


Figure 6. Hybridization and ligation-based detection of SNP-mtDNA.

Table 3. Sequences used in the hybridization-based mtDNA assay*

Sequence Name	Nucleotide Sequence
MT-CYB 14766 C>T Target	5'-ATGACCCCAATACGCAAAATTAACCCCCTAATAAAATTA-3'
MT-CYB SNP 14766 C>T Capture	5'-/AmMC12/TAATTTTATTAGGGGGTTAA-3'
MT-CBB 14766 C>T Detection	5'-/5Phos/TTTTGCGTATTGGGGTCAT/3Bio/-3'

* AmMC12 indicates amine modification with a C12 chain for attachment chemistry, and Phos refers to a phosphate modification to facilitate future ligation between a phosphate group and a neighboring hydroxyl group.

These sequences were tested in a sandwich hybridization protocol without the ligation, which yielded a LOD of 16 fM. This approach afforded femtomolar sensitivity, but no specificity to distinguish between the wild type and SNP sequences. Results did not improve with the incorporation of a ligase. Any future approaches may include the use of modified nucleic acids to enhance the specificity of the assay.

1c. Screen blood samples from 'Human-In-Mouse' (HIM) Model obtained from Kuperwasser (Months 12-24). HIM tumors will be created from human breast epithelial cells collected from discarded tissues of women who have undergone reduction mammoplasty surgeries. For more information/details on model: *Nature Protocols*, 2006, 1, 595-599.
This task is essentially completed. We successfully created a mouse model using low inoculums of LNCaP cells to study PSA in mouse serum over time. Through this study, we demonstrated that SiMoA can be used to measure biomarkers in serum prior to tumor formation, and the data has been published. Since this study proved that our technique is valid for the early detection of tumors, we chose to pursue the study of clinical samples in lieu of developing more mouse models since this will yield more meaningful data as an end result.

Task 2. Apply single molecule diagnostic technique (developed in Task 1) to analysis of human serum samples (Years 3-5).

2a. Screen clinical samples obtained from Buchsbaum for presence of the markers that could be detected in HIM model (Months 25-30).

Dr. Buchsbaum is consenting patients and collecting clinical samples. These samples will be tested against all available SiMoA assays as they become available. Preliminary experiments have been performed testing human samples using commercially available sources as detailed below in section 2b.

2b. Iterate 2a-2d until a sufficient number of markers has been identified for further investigation. Some candidate markers are expected to either not be found in the clinical samples or will not be predictive of disease state. Other markers may be difficult to measure due to lack of suitable binding reagents. As these markers drop out, others will be added to ensure there is a sufficient number when sample set is expanded (2c) (Months 25-48).

Breast cancer serum samples and healthy controls were obtained from the commercial supplier BioreclamationIVT. These samples were tested with our developed SiMoA protein assays to quantify the levels of candidate biomarkers in clinical samples and validate the usefulness of each biomarker. **Figure 7** shows a graph of different markers that have been measured in healthy serum and breast cancer serum. Section 1b details efforts to multiplex these assays. We are currently testing these multiplex assays for cross-reactivity and sensitivity.

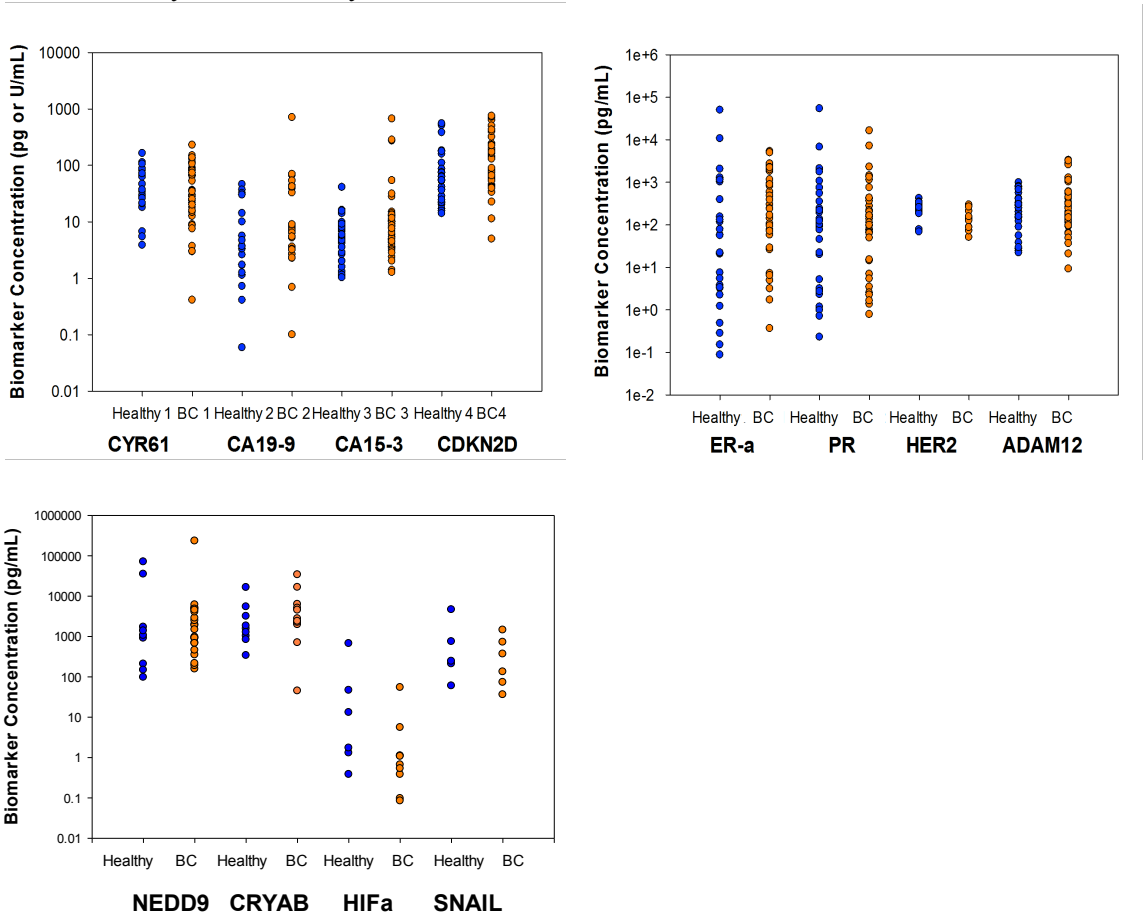


Figure 7. Measurements of protein biomarkers in commercially available healthy and breast cancer serum.

There have been many reports showing that extracellular vesicles carry molecular information, which can affect neighboring cell signaling pathways as they are released. In addition to testing untreated serum samples for the presence of NEDD9 and CRYAB, non-denaturing lysis buffer was added in the serum samples and levels of both NEDD9 and CRAYB were then subsequently evaluated with hopes of finding these proteins within extracellular vesicles such as exosomes in the serum. Interestingly, we found substantial NEDD9 and CRYAB levels in the serum samples treated with lysis buffer in both healthy and

breast cancer samples from which both protein levels were undetectable in the serum alone (Figure 5A, B). The SiMoA calibration curve was not affected upon treatment with lysis buffer confirming the lysis buffer did not impact nonspecific binding. This data is a proof of concept study that demonstrates that secreted candidate biomarkers can be detected using SiMoA after adding lysis buffer to the serum samples. In the future, more samples will be tested with the addition of lysis buffer and the levels of the candidate biomarkers will be evaluated to correlate with disease outcomes.

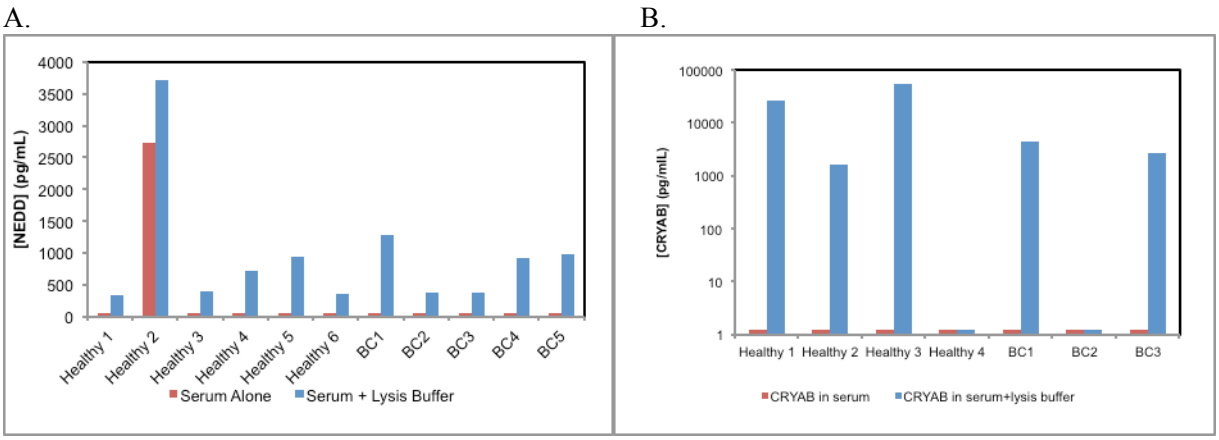


Figure 8. A) NEDD9 levels. B) CRYAB in healthy and breast cancer serum samples treated with the lysis buffer.

Exosomes were collected from serum samples with an average size of approximately 110 nm determined by Nanosight particle tracking system. **Figure 9** shows a comparison of ADAM12 and NEDD9 levels in serum versus the exosome lysate. We were able to find ADAM12 in the exosome lysate; however, the amount of protein released from the exosomes was much lower than in the original serum in all but one of eight samples. Unlike ADAM12, NEDD9 is not accessible for detection in serum. In fact, NEDD9 was only measured in one serum sample. However, NEDD9 was measured in serum samples following a lysis protocol for the serum, as described above. This indicates that NEDD9 may be contained in exosomes in the serum and is released upon lysis. We were able to detect NEDD9 in the exosome lysate in 3 out of 8 samples, but the amount of NEDD9 detected from the exosome lysate does not correlate well with the amount measured in serum or serum lysate. This result suggests that exosomes are not the only kind of vesicles that may contain NEDD9. In comparison to ADAM12 and NEDD9, we could not detect HER2 in the exosome lysate, even though the serum levels of HER2 are high for all the samples tested (average 585 pg/mL).

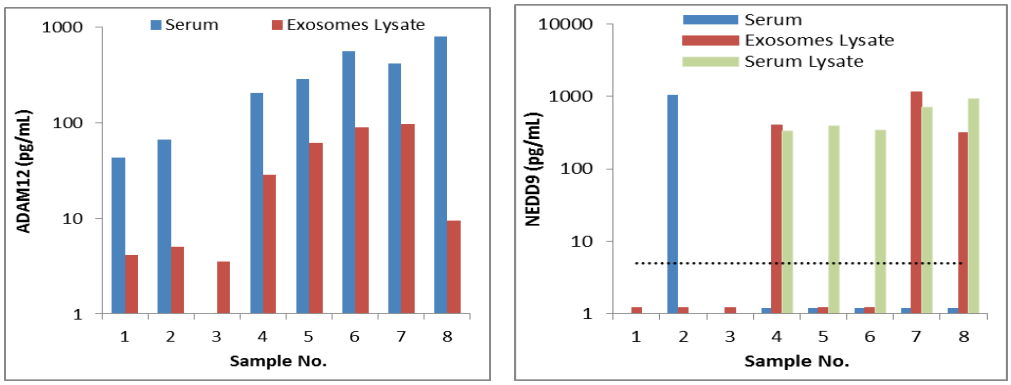


Figure 9. ADAM12 and NEDD9 protein level comparison between exosome lysate and original serum.

Based on the results above, we have determined that proteins inside exosomes are not good biomarker candidates for use with the ultrasensitive SiMoA platform. The proteins of interest were either measurable

directly in serum (e.g., HER2 & ADAM12) with higher level than in exosomes, or could be detected by lysing the serum directly (e.g., NEDD9). Overall, the results here demonstrate that exosomes are not superior to serum (or serum lysate) as a clinical sample for ultrasensitive protein marker detection. Further investigation of exosomal content as biomarkers may shift from protein to nucleic acids like microRNAs.

2c. Expand sample set and run clinical samples with all assays developed for all markers (Year 5).
These studies will be initiated with the expanded sample set provided by Dr. Buchsbaum.

2d. Determine which markers in blood correlate with disease using data processing and computational methods available in the laboratory (Year 5).
Preliminary multivariate data analysis has been carried out to assess the predictive capabilities of several protein biomarkers as detailed in Task 2b of the April 2015 Quarterly report. The supervised multivariate statistical method of partial least squares-discriminant analysis (PLS-DA) was employed to assess PR, ERα, and CDKN2D. Results were promising with one model yielding 84% accuracy in distinguishing healthy individuals from early stage (I-II) breast cancer patients (**Table 4**). Additional models will be made with the remaining protein biomarkers. It is important to note that the samples used as input into the model are not ideal since all the BC samples had undergone some treatment (surgery, chemotherapy, steroid therapy) prior to sampling, thereby confounding the results since some of them may have returned to a more "normal" protein biomarker phenotype. On the other hand, these models will be informative for the analysis of the prospective clinical samples being collected by Dr. Buchsbaum.

	Description	Precision	True Positives	AUC	Number of samples	Overall Accuracy
Model 1	Healthy	72%	90%	0.88	31	81%
	BC Stage I-IV	92%	75%	0.88	44	
Model 2	Healthy	82%	87%	0.91	31	84%
	BC Stage I-II	86%	80%	0.91	30	
Model 3	Healthy	86%	81	0.88	31	45%
	BC Stage I-II	62%	70%	0.78	30	
	BC Stage III-IV	0%	0%	0.63	14	
Model 4	TNBC	50%	78%	0.73	14	68%
	Luminal	86%	63%	0.73	30	

Table 4. PLS-DA analysis results for different populations of breast cancer serum samples.

Task 3. Develop single cell analysis methods to determine composition of a primary tumor.

3a. Select approximately ten candidate markers (SNPs, proteins, etc.) of value for single breast cancer cell analysis in conjunction with collaborators (Months 1-9).
We decided to utilize the pre-established PSA assay to validate SiMoA at the single cell level. This assay has proven to be extremely sensitive and our goal is to take advantage of the sensitivity of this established assay to establish the feasibility of performing single cell protein analysis.[1] In order to test this assay with cells, we have cultured the prostate cancer cell line LNCaP (cell lines were purchased from ATCC and donated by the Kuperwasser lab). We are also working towards developing single cell assays using the HER2 and PR SiMoA assays established in Task 1. Breast cancer cell lines T-47D, BT-474, MDA MB231 and MCF7 will be used to quantify these markers in single cells.

3b. Detect presence of mtDNA by screening for many sequences using cultured cell lines and HIM tissue samples to select which markers to use. Will require development of mtDNA microarrays and sample screening (Months 6-18).
The selection of mtDNA markers is being performed in conjunction with the Sonenshein laboratory (see Sonenshein Task 1 below). The DNA assay developed above in Task 2 will be the basis for the assays we plan to use for Tasks 3b and 3c.

3c. Develop assays for the selected markers (Months 6-24).
The SiMoA assay for PSA has already been developed. PSA was used as a proof of concept assay because it is well established. SiMoA assays for PR, ER-α, and HER2, and others have been developed as described in Task 1. These markers will be useful for single cell studies of breast cancer cell lines such as T-47D, BT-474, MDA MB231 and MCF7.

3d. Develop single cell assays for selected markers (Months 18-36).
We previously demonstrated the ability to detect PSA in single LNCaP cells, and have expanded that work to characterize breast cancer cell lines of different subtypes. The goal of this work is to use single cell SiMoA to classify and identify breast cancer cell subtypes in a heterogeneous tumor population with single cell resolution. We have cultured three different breast cancer cell lines to study the distribution of protein expression based on breast cancer subtype. MDA-MD-231, BT-474, and MCF7 cell lines were selected to represent triple negative breast cancer (TNBC), luminal B, and luminal A subtypes, respectively. In the past year, we have measured the levels of eight different protein biomarkers in bulk cell populations containing roughly 100 cells for each cell line (**Table 5**). These experiments will inform how to proceed to multiplexing markers for single cell studies.

The results in **Table 5** clearly demonstrate the differences in protein expression observed between different cell lines. BT-474 overexpresses HER2 and PR, and shows measurable expression of ERα, CYR61, CA15-3, CA19-9, and ADAM12. MCF7 cells show considerably lower expression of HER2, which is consistent with the luminal A subtype. Likewise, MDA-MB-231 cells show higher concentrations of CYR61 and no detectable levels of HER2, PR, or ERα, as expected for TNBC.

Table 5. Comparison of Breast Cancer Protein Expression in ~100 Cells

Marker	BT-474	MCF7	MDA-MB-231	Assay LOD
HER2	77	1.2	Below LOD	0.3
PR	9.2	Below LOD	Below LOD	0.2
ER-α	3.7	1.7	Below LOD	0.2
CYR61	2.5	0.54	7.2	0.03
CA 15-3	0.21	Below LOD	Below LOD	0.1
CA 19-9	0.021	--	--	0.003
ADAM12	9.2	3.8	4.8	3
HIF1α	0.083	0.031*	0.04	0.005

* Concentration tested in a 50-cell dilution

We have also quantified HER2 and CYR61 protein molecules in single BT-474 and MDA MB231 cells, with results shown in **Figure 10**.

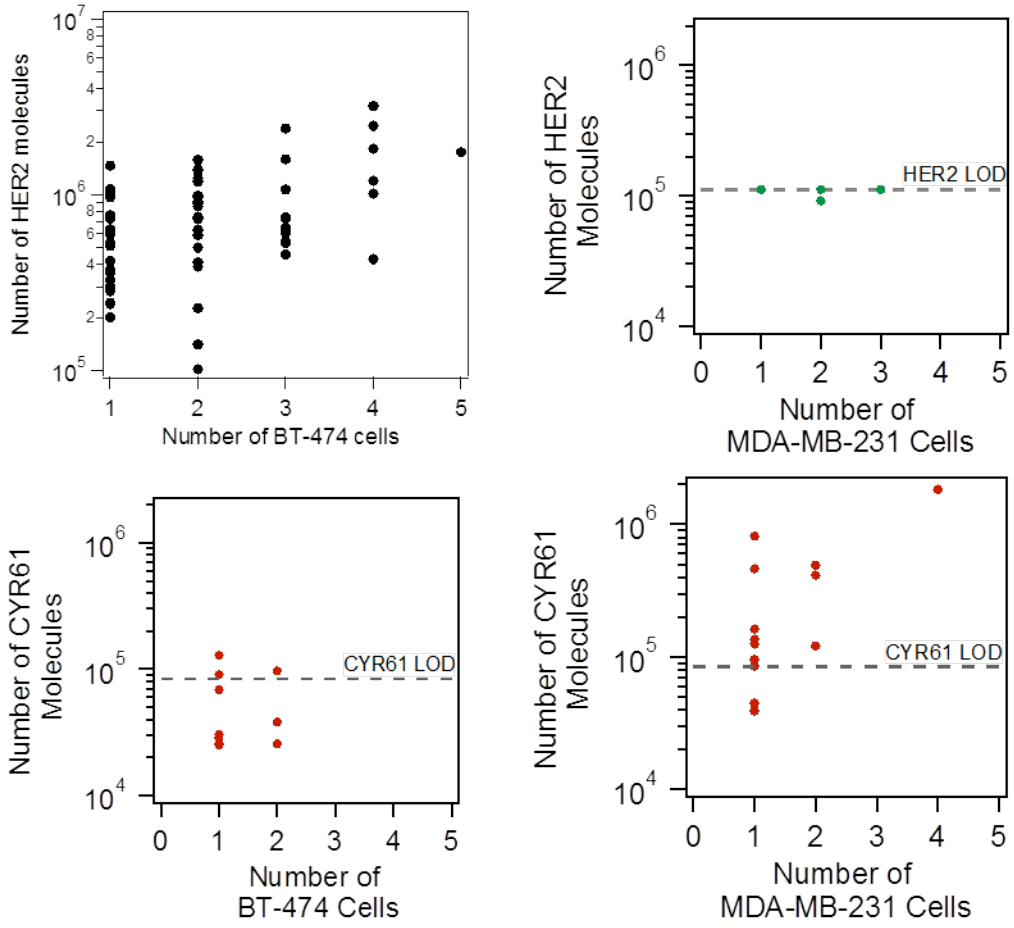


Figure 10. Number of HER2 and CYR61 molecules detected in isolated BT-474 and MDA MB231 cells.

3e. Obtain tissue samples from HIM (Kuperwasser) and perform single cell analysis (Months 36-48).

These studies have not been initiated per SOW.

3f. Determine if one can observe rare cells with more aggressive genotypes or protein levels in HIM samples (Months 42-54).

These studies have not been initiated per SOW.

3g. Confirm presence of the single cells in human breast cancer biopsy samples obtained from Buchsbaum (Months 48-60).

These studies have not been initiated per SOW.

Daniel T. Chiu, PhD, *University of Washington, Seattle Campus Box 351700 Seattle, WA 98195-1700*

Task 1. Work with Walt lab to develop/refine single-molecule and single-cell techniques for analyzing protein and cell biomarkers (Months 1-36).

Task 1 *will use cultured cell lines, commercially available serum samples, and serum samples from HIM model.*

For Task 1, we have developed a new technology for the detection and analysis of cell biomarkers, in particular, circulating tumor cells (CTCs). CTCs have emerged as an important and valuable biomarker for the prognosis of breast cancer, and the sensitivity we have demonstrated makes CTCs a robust biomarker for prognosis and as an indicator of treatment efficacy. The high sensitivity also has the potential to make CTCs a diagnostic tool if they are present at sufficient levels in blood. Below, we describe each subtask in more detail.

***1a. Develop microfluidic devices for sample preparation and optical manipulation. Initial devices will contain only sample preparation module for use by Walt lab (Years 1-2*).**
We have completed this task, where we now have a new technology called eDAR (ensemble Decision Aliquot Ranking) with which we can isolate circulating breast cancer cells from the periphery blood of patients with over 90-95% recovery efficiency in a background of over billions of blood cells in less than 20 minutes. We continue to make improvements and refinements to this new technology. One example is the use of better fluorescent probes for targeting breast cancer cells so we do not miss sub-types of cancer cells from patients, given the highly heterogeneous nature of these cells. Another example is the downstream genomic and proteomic analysis of the isolated circulating breast cancer cells.

1b. Integrate assays and optical methods developed in Walt lab into a microfluidic device for the analysis and profiling of biomarkers (Years 1-2).
We have completed this task and now have a low throughput method for the isolation and analysis of individual CTCs (circulating tumor cells), including performing digital PCR and SiMoA on the isolated single cells. This low throughput method relies on manual selection and transfer of cells under a microscope. To improve the throughput of this method, we are currently finalizing a sequential eDAR and single-cell dispensing system so that each cell isolated by eDAR can be deposited into a well of a 96 well plate (more details below under 2d). This capability also will be useful for automated deposition of single cells into a 96-well plate format for use with SiMoA.

1c. Develop integrated microfluidic and optical techniques for single cell analysis using model cell lines (Years 1-3).
We have completed this task and have developed several single-cell analysis techniques for interrogating the isolated CTCs, especially for the quantification of gene and protein expression. As described in previous reports, the single-cell analysis we demonstrated on breast cancer cells includes the imaging of a large panel of protein biomarkers using cycles of sequential protein labeling followed by fluorescence imaging and photobleaching. We also carried out digital PCR quantification of mRNAs in individual breast cancer cells. Finally, we performed both functional assays on the isolated breast cancer cells as well as the culture and expansion of single breast cancer cells. To further improve the analysis throughput, we have also explored and reported the parallel manipulations of single cells using bipolar electrode (BPE) based dielectrophoresis (DEP).

Briefly, for the BPE-DEP based parallel manipulations of single cells, we have developed the theory that describes this technique and also carried out the first demonstration. The first paper that presented this work was published this year. **Figures 11** and **12** show the use of BPE-DEP for the manipulation of single

cells in a microfluidic channel. Figure 11 shows the use of BPE-DEP to repel single cells in flow, and Figure 12 shows the trapping (left panel) and release (right panel) of single cells.

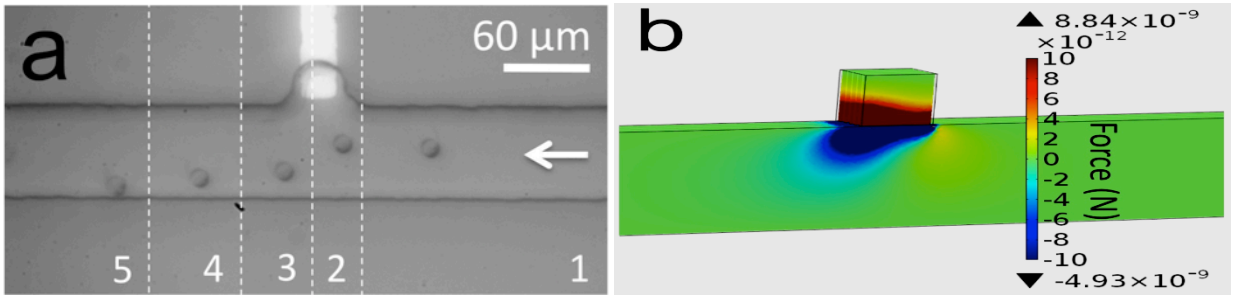


Figure 11. A) A series of optical micrographs, which show nDEP repulsion of a single B cell from the BPE tip under AC-only electric field in Tris DEP buffer. Each image slice (numbered sequentially 1-5) is separated by 2.5 s. $E_{RMS,avg}$ = 5.7 kV/m (t = 0 s) to 17.7 kV/m (t = 5 s). B) Simulated magnitude of the y-component of F_{DEP} in the xy-plane at z = 5 µm.

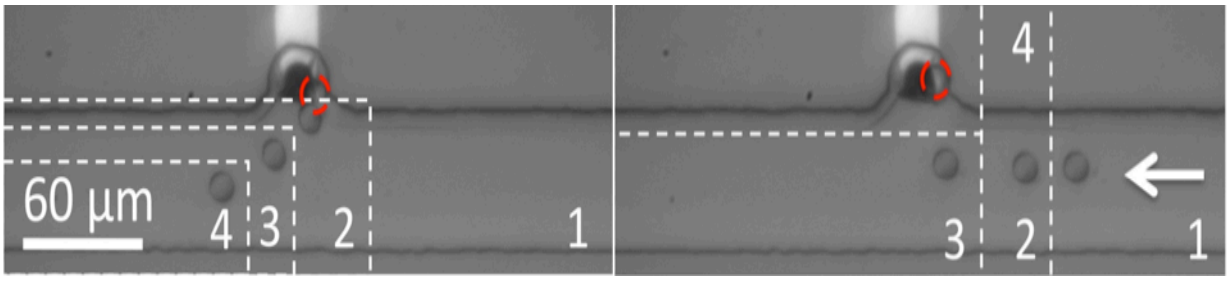


Figure 12. Left panel: nDEP attraction of a B-cell toward the BPE cathode in phosphate DEP buffer (4 s/slice). $E_{DC,avg}$ = 1.5 kV/m, $E_{RMS,avg}$ increased from 5.7 kV/m to 28.3 kV/m from t = 0 s (slice 1) to t = 8 s (slice 3). Right panel: Release of the trapped cell (2 s/slice) from (d) upon subsequent decrease of $E_{RMS,avg}$ to 5.7 kV/m (from slice 1 to slice 2).

Task 2. Work with the Walt lab to apply these methods to breast-cancer patient samples for the detection and validation of protein and cell biomarkers *(Months 25-60).*

2a. Develop/submit amended proposal to University of Washington IRB to permit secondary use of currently archived patient samples (Months 1-12).

This task has been completed per previous reports.

2b. Apply sensitive techniques (1a and 1b) for the retrospective analysis and validation of biomarkers from archived patient samples (Months 25-60).

We propose to use archived blood plasma samples from breast cancer patients. One hundred archived blood plasma samples already exist in the Chiu lab from a previous three-year study on detection of biomarkers in breast cancer patients. We are making good progress on this task and we have sent about 50 samples to the Walt lab for SiMoA analysis.

2c. Apply sensitive techniques (1a and 1b) for the prospective analysis and validation of biomarkers from patient samples (Months 36-60).

We are making good progress on this task and have analyzed 25 patient samples thus far. **Table 6** below summarizes the results.

Table 6.

Sample Code	Labeling strategy	CTC in 2 ml	Subtype	Stage
B-140000007	1 step EpCAM	8	Triple -	Stage 4
B-140000053	1 step EpCAM	1.5	ER+, PR- Her2+	Stage 2
B-140000089	1 step EpCAM	0	ER/PR+ Her2-	Stage 2
B-140000122	1 step EpCAM	4	ER/PR - Her2+	Stage 4
B-140000169	Cocktail (EpCAM, HER2, EGFR, N-cadherin)	32	ER/PR+ Her2-	Stage 3
B-140000169	1 step EpCAM	8	ER/PR+ Her2-	Stage 3
B-140000477	Cocktail (EpCAM, HER2, EGFR, N-cadherin)	15	ER/PR + Her2 -	Stage 2
B-140000865	1 step	0	ER/PR- Her2-	Stage 4
B-140000931	1 step	29*	ER/PR+ Her2-	Stage 4
B-140002443	Cocktail(EpCAM, HER2, EGFR, N-cadherin)	19	ER/PR+ Her2-	Stage 4
B-140001155	Cocktail (EpCAM, HER2, EGFR, N-cadherin)	27*	ER+PR-Her2-	Stage 1A (T1N0)
B-140001155	1 step	5*	ER+PR-Her2-	Stage 1A (T1N0)
B-140001165	1 step	4*	ER/PR+ Her2-	Stage 2
B-140000068	1 step	48*	ER/PR+ Her2-	Stage 2
B-140001108	1 step	1, 4*	ER/PR+ Her2-	Stage 2
B-140001146	1 step	1, 5*	ER/PR- Her2-	Stage 2
B-140001146	Cocktail (EpCAM, HER2, EGFR, N-cadherin)	1	ER/PR- Her2-	Stage 2
B-140001196	1 step EpCAM	44	ER+ PR- Her2+	Stage 4
B-140001196	Cocktail (EpCAM, HER2, EGFR, N-cadherin)	75	ER+ PR- Her2+	Stage 4
B-140001222	Cocktail (EpCAM, HER2, EGFR, N-cadherin)	57	ER+ PR+ Her2-	Stage 2
B-140001222	1 step EpCAM	31	ER+ PR+ Her2-	Stage 2
B-140001362	1 step EpCAM	3	ER- PR- Her2+	Stage 4
B-140001385	1 step EpCAM	0	ER+PR+ Her2-	Stage 2
B-150001464	1 step EpCAM	2	ER/PR+ Her2-	Stage 2
B-150001464	Cocktail (EpCAM, HER2, EGFR, N-cadherin)	18	ER/PR+ Her2-	Stage 2

*These CTCs also expressed CD45

Based on the above clinical studies, we found CTCs isolated from patients to be highly heterogeneous, both in terms of the amount of cell-surface proteins that the CTCs express and the types of cell surface proteins present on the CTCs. To increase the number of CTCs isolated from patient samples, especially from early stage patients, we have carried out the following three studies: (1) A systematic study to determine the minimum number of proteins needed on CTCs in order for eDAR to capture them efficiently; (2) develop ultra-bright probes based on semiconducting polymer dots for the detection of cells with low biomarker expression; (3) improve the ability of eDAR to capture CTCs with variable patterns of biomarker expression by developing a cocktail of antibodies to label multiple proteins on CTCs. We describe each of these developments below.

(1) Limit of detection (LOD) of eDAR. eDAR requires the use of fluorescent antibodies that bind to the biomarkers on the cancer cells to "light up" the cancer cells so they can be located and isolated. If there is

a low level of biomarkers on the CTC, a correspondingly small number of fluorescent antibodies are bound to the CTC and the cell may not be sufficiently bright for eDAR to pick up. To understand the LOD of eDAR, we characterized the brightness of the different breast cancer cell lines known to express various amounts of the cell surface protein EpCAM (currently the cell surface marker used for isolation of CTCs) and then determined the number of EpCAMs present on each cell based on the number of fluorescent antibodies bound to EpCAM. Similarly, using MDA-MB-231 cells (which mimic triple-negative breast cancer), we have determined the number of several cell-surface biomarkers specific to breast-cancer cells. **Figure 13** shows the results; here, the number of cell-surface biomarkers was measured based on PE (phycoerythrin) fluorescence where there was a one-to-one correlation between the number of PE and the number of cell-surface biomarkers.

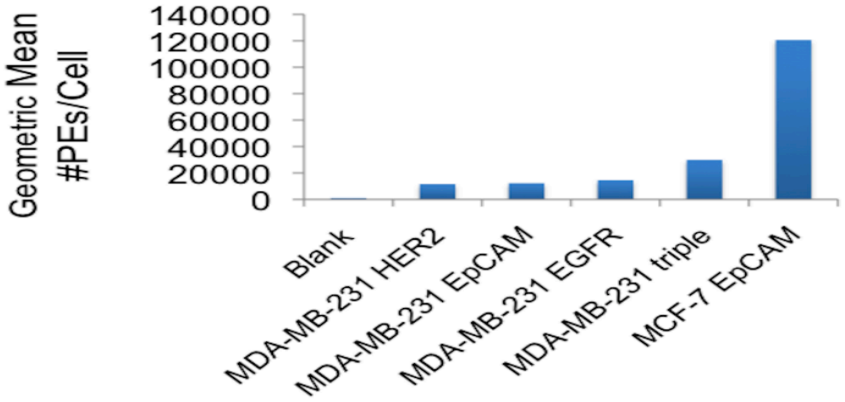


Figure 13. MDA-MB-231 cells labeled with three antibodies (EpCAM, HER2, EGFR) and a cocktail of all three, demonstrating that the antibody brightness is additive.

From these experiments, we found our current eDAR LOD is 5,800 EpCAMs and LOQ (limit of quantitation) is 23,500 EpCAMs per cell. While LOD is defined as 3 times the standard deviation (σ) of the background, LOQ is defined as 10σ, which is the threshold eDAR uses for robust sorting and to minimize any potential false sorting caused by fluctuations in the background signal.

(2) Ultra-bright probes for the isolation of cells with low biomarker expression using eDAR. CTCs with low protein expression are generally categorized as those with less than 10,000 molecules/cell. While our LOD as described above is below this level, our LOQ is above this threshold. To address this issue, we tried to improve our LOQ by using a two step labeling protocol; at present, we use fluorescently labeled primary antibodies, but the use of fluorescently labeled secondary antibodies may amplify our signal two fold because of multiple binding of secondary antibodies to the primary. This strategy was partially successful, but to significantly increase the signal, we have also developed a new fluorescent probe based on semiconducting polymers to enhance the fluorescence signal of bound antibodies. This new probe is an important step towards addressing this issue of low biomarker expression because even if a small number of antibodies were bound to the CTC, the brightness of the probe would compensate for the low level of biomarker expression to light up the CTC.

Semiconducting polymer dots (Pdots) represent a new class of ultra-bright fluorescent probes for biological detection and imaging. Their diameters can be tuned from just a few nanometers to tens of nanometers. The motivation for adapting fluorescent semiconducting polymers into nanoparticle labels stems from a number of favorable characteristics, such as their large absorption coefficients, high quantum yields, fast emission rates, and excellent photostability. The resulting Pdots exhibit extraordinarily high fluorescence brightness under both one-photon and two-photon excitations, a factor of 10^2-10^4 higher than conventional fluorescent dyes, and a factor of 10-10^3 higher than even Qdots. **Figure 14** shows the brightness of Pdots.

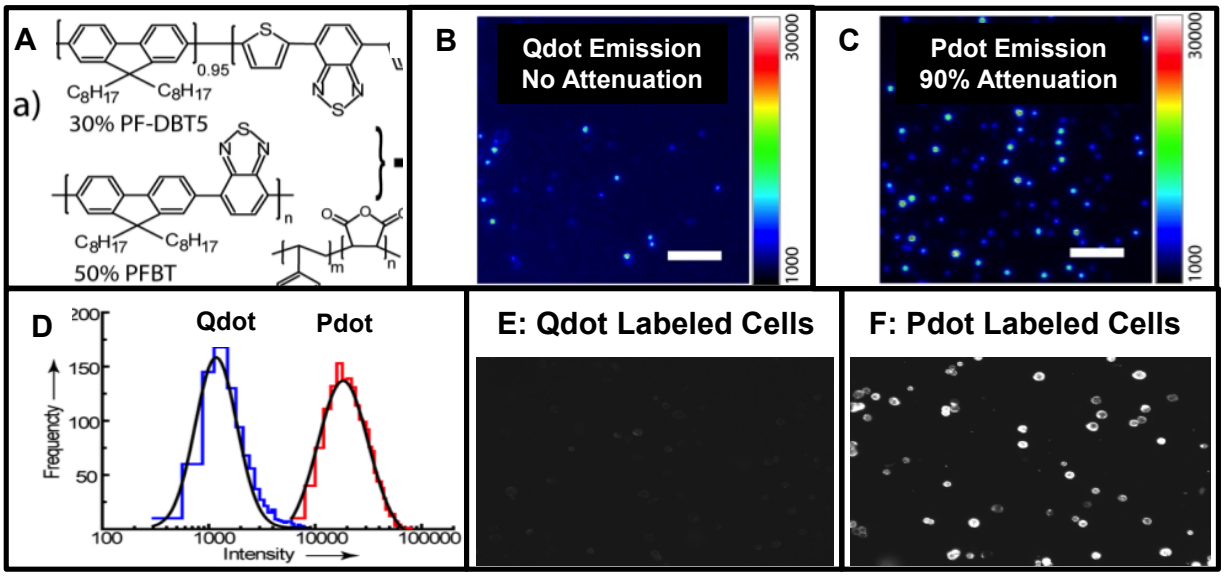


Figure 14. Extraordinary brightness of Pdots. **A**) Absorption (Abs; green curve) and Emission (Em; red curve) spectra of an example Pdot. (**B - D**) Brightness comparison between Qdots & Pdots: **B**) A single-particle fluorescence image of Qdot. **C**) A single-particle fluorescence image of Pdot under identical imaging conditions as in B) but after 90% of emitted light was attenuated by a neutral-density filter placed in front of the camera. **D**) A plot of the single-particle fluorescence intensity distribution of Qdots (blue histogram) and Pdots (red histogram). **E, F**) Fluorescence images showing Pdot labeled cells (F) are much brighter than Qdot labeled cells (E) under identical experimental conditions.

Figure 14 A-D shows a red emitting (~ 640nm emission; ~ 470nm excitation) Pdot; these Pdots are much brighter than comparable red emitting Qdots (Fig. 14B & 14C). Note here that Pdot emission was attenuated by 90% in Fig 14C. Pdot labeled cells are also much brighter than Qdot-tagged cells (Fig 14E); under imaging settings where Pdot-tagged cells did not saturate the camera, Qdot labeled cells were barely visible. This result was further confirmed with eDAR by comparing the fluorescence intensity of cells labeled with Qdots versus Pdots under identical experimental conditions. In fact, using single-particle imaging, we have routinely collected 10^9 photons from single Pdots, which is 1,000× to 100,000× more than many single dyes and green fluorescent proteins (GFPs), and 10-100 times more than Qdots.

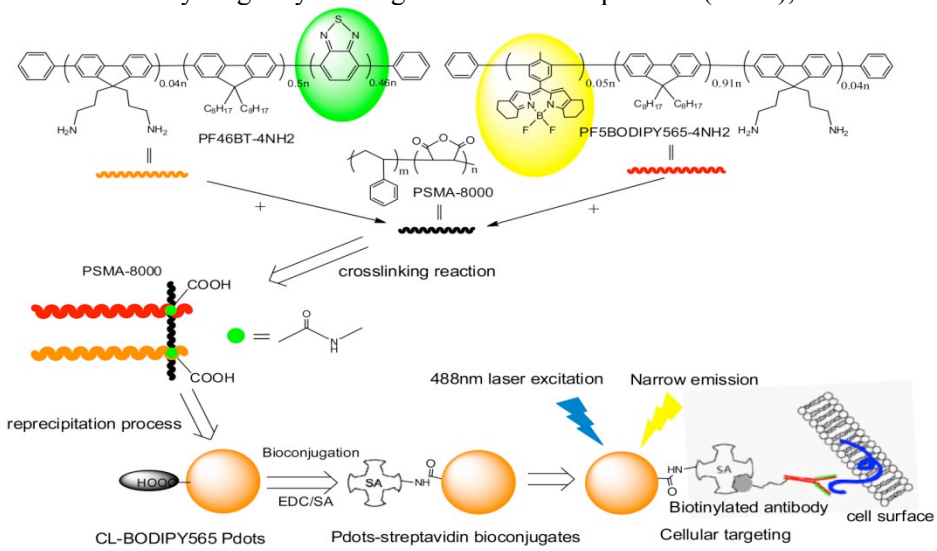


Figure 15. Schematic illustration of small, bright and narrow emissive crosslinked semiconducting polymers and Pdot-bioconjugates for cellular labeling and use in eDAR.

However, the Pdots so far have broad emission and thus are not suitable for use with eDAR, which requires narrow-emission Pdots for multiplexed detection of several biomarkers. Additionally, there are no suitable Pdots that emit in the same wavelength range as PE, which is the current probe we use for eDAR. As a result, we have developed a new Pdot with narrow emission and a spectral profile (yellow emission) that is similar to PE for use with eDAR (Figures 15-17). The new yellow Pdots consist of donor and acceptor polymers, which were together cross-linked with one amphiphilic functional polymer (**Figure 15**). In detail, a narrow emissive fluorescent dye, boron-dipyrromethene (BODIPY) was copolymerized into the polymer as the acceptor molecule, poly[9,9-dioctylfluorenyl-2,7-diyl)-co-1,4-benzo-{2,1'-3}-thiadiazole)] (PFBT) was used as the donor molecule. More importantly, the emission spectrum of the donor has a very good overlap with the absorption of the acceptor to enable efficient FRET between them. We then developed a crosslinking strategy to covalently bond these fluorescent semiconducting polymers together with the amphiphilic functional polymer, poly(styrene-co-maleic anhydride) (PSMA), and successfully obtained crosslinked polymers with the donor and the acceptor parts. Small and stable Pdots were formed via the nanoprecipatation process, and they gave high fluorescence brightness and narrow emission by the FRET mechanism.

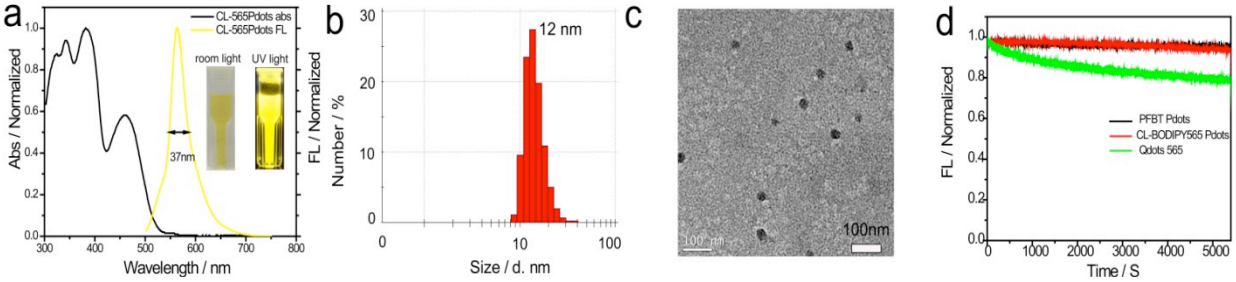


Figure 16. A) Absorption and Fluorescence spectra of CL-BODIPY 565 Pdots in water. B) The histograms of the distribution of the sizes of CL-BODIPY 565 Pdots, measured by DLS (the mean size is 12 nm). C) TEM images of CL-BODIPY 565 Pdots. D) photobleaching curves of PFBT Pdots (black line), CL-BODIPY 565 Pdots (red line) and Qdots-565 (green line).

The crosslinked Pdots (CL-BODIPY 565 Pdots) show narrow emission centered at 564 nm with a FWHM as narrow as 37 nm (**Figure 16A**). To the best of our knowledge, this is the narrowest emission bandwidth among various Pdots reported so far. To evaluate its photophysical properties, the Pdots were prepared by the nanoprecipatation method with an average particle size of 12 nm, as characterized by dynamic light scattering (DLS) and transmission electron microscopy (TEM) (Fig. 16B & 16C). Commercially available Qdots 565 were dispersed in MilliQ water and measured by DLS under the same conditions and found to have a comparable particle size of ~13 nm. Our developed CL-BODIPY 565 Pdots also exhibited remarkable photostability (measured on the fluorimeter equipped with 450 W Xenon lamp, and the excitation wavelength was centered at 488 nm). The photobleaching process has been kept for up to 1.5 hours, which did not result in observable decrease in fluorescence intensity of CL-BODIPY 565 Pdots (Fig. 16D), while Qdots 565 exhibited an obvious single exponential photobleaching decay. The comparable high photostability of CL-BODIPY565 Pdots to our previous PFBT Pdots should be attributed to the incorporation of PFBT donor polymer into the CL-BODIPY 565 Pdots.

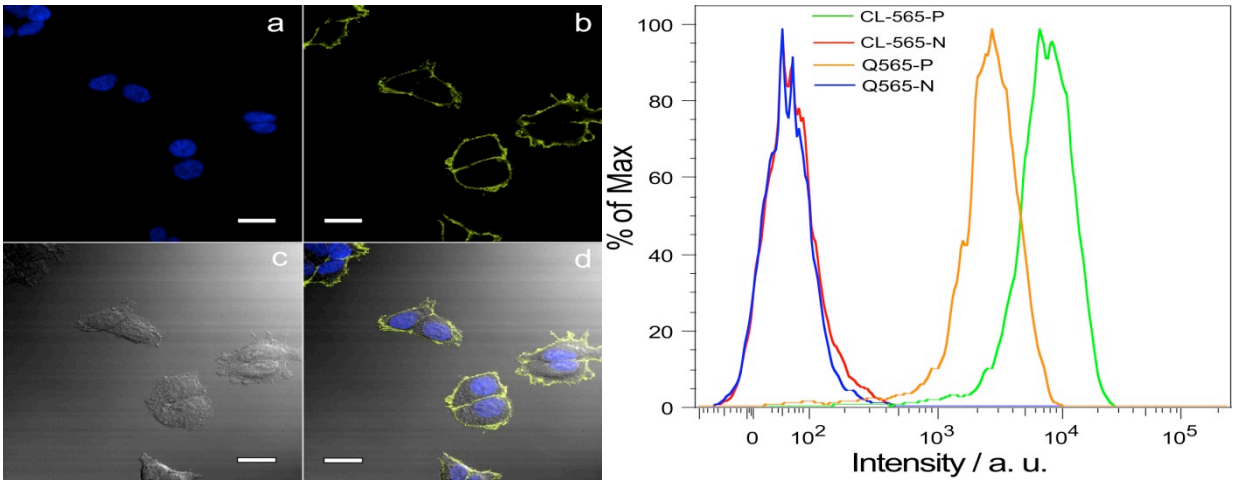
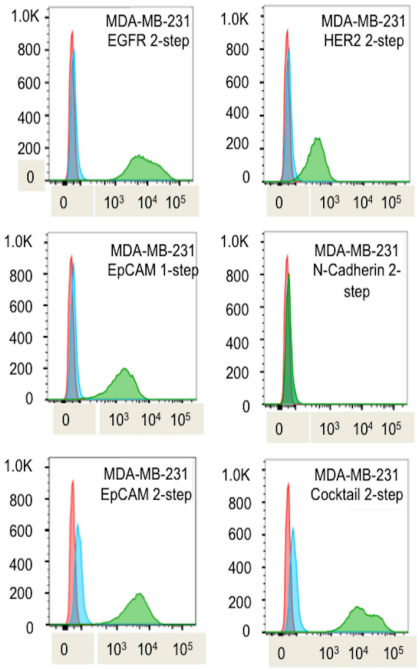


Figure 17. (A-D) Confocal fluorescence microscopy images of MCF-7 breast cancer cells labeled with Pdot-streptavidin against biotinylated anti-EpCAM: A) blue fluorescence from the nuclear stain Hoechst 34580. B) Yellow fluorescence from Pdot. C) Nomarski (DIC). D) Combined DIC and fluorescence images. Scale bars: 20 μm. E) Flow-cytometry measurements of the intensity distributions of MCF-7 breast-cancer cells labeled via non-specific binding (N: negative control) and positive specific targeting (P: positive control) using Qdots 565 (Q565-N: Qdot negative control; Q565-P: Qdot positive control), Pdots (CL-565-N: CL-BODIPY 565 negative control; CL-565-P: CL-BODIPY 565 positive control). All Qdots and Pdots were conjugated with streptavidin.

To evaluate the performance of the CL-BODIPY 565 Pdots for biological applications, they were applied in flow cytometry and cellular imaging by labeling with the breast cancer MCF-7 cells. Pdots were conjugated with streptavidin via the 1-ethyl-3-[3-dimethylaminopropyl]carbodiimide hydrochloride (EDC) catalyzed coupling reaction. Further, these Pdot-bioconjugates were used to label MCF-7 cells, which were first incubated with biotinylated anti-EpCAM antibodies. **Figure 17** shows the results. The specific cellular labeling of the Pdot-SA probes was confirmed by confocal fluorescence imaging (Fig. 15A-15D). From the results, it is evident Pdot-SA probes were effectively labeled with biotin anti-EpCAM receptors on the MCF-7 cell surface; however, no fluorescence was detected in the negative control experiment which was carried out in the absence of the biotinylated primary antibody. Therefore the result proved that Pdot- SA could be used as probes for specific cellular labeling without non-specific binding. Figure 17E shows the flow cytometry results. There was excellent separation between Pdot-streptavidin bioconjugates labeled cells and the negative control where cells were incubated with Pdot-streptavidin bioconjugates but in the absence of the biotinylated primary antibody. MCF-7 cells labeled with Pdot-streptavidin bioconjugates (CL-BODIPY565 Pdot-SA) exhibited about 5 times higher intensity than that labeled with Qdot565-streptavidin bioconjugates under identical experimental conditions.

(3) Improve the ability of eDAR to capture CTCs with variable patterns of biomarker expression by developing a cocktail of antibodies to label multiple proteins on CTCs. Many CTC detection methods, including ours, take advantage of the epithelial origin of CTCs, which provides the cells with surface markers that are distinct from those on other cells in the blood. However, for a tumor cell to migrate into the bloodstream, it is reported to undergo an epithelial-to-mesenchymal transition (EMT), losing some of its epithelial characteristics, such as structural rigidity, cell adhesion and epithelial markers, such as EpCAM and cytokeratin, and taking on a more mesenchymal phenotype. The phenotypic changes undergone by CTCs during the metastatic process mean that these cells may not express high levels of the epithelial markers, such as EpCAM, that are targeted by current CTC detection techniques. These "EpCAM^{low}" CTCs recently have received attention because they potentially can evade detection and have been linked with enhanced invasiveness and migration. Accurate population statistics describing CTC expression of epithelial markers is unavailable. More importantly, immunoaffinity techniques

traditionally quantify recovery rate (percentage of cells captured) using cultured cancer cells that are EpCAM^{high} (e.g. the breast cancer cell line MCF-7) thus leading to an *upper estimate* of CTC recovery.



Cell Line	Geometric Mean PEs
MCF7:	
1-step EpCAM	353205
Control MCF7 PE-IgG1	*41*
2-step	590456
Control IgG2	*364*
MDA MB 231:	
2 step EGFR	30233
2-step HER2	1484
2-step N-Cad	58
Control IgG1	*56*
2-step Cocktail	38638
Control IgG1&2	*171*
1-step EpCAM	5890
Control PE-IgG1	*69*
2-step EpCAM	14873
Control IgG2	*182*

Positive Cell Population
IgG Control
Blank

Figure 18. Development of an antibody cocktail labeling strategy, where we tested the specificity of 4 different primary antibodies (EpCAM, N-Cadherin, Her2, EGFR) against breast cancer cells. As a model, we used MDA-MB-231 cells, which mimic triple-negative breast cancer and which have low EpCAM expression. For this experiment, separate populations of MDA-MB-231 cells were labeled with each primary antibody and the corresponding IgG antibody followed by the secondary antibody to verify specificity of primary antibody labeling.

To address this issue, we used a cocktail of CTC-specific primary antibodies targeting both mesenchymal and epithelial cell surface markers as well as a single dye-linked secondary antibody. The use of a single fluorescently labeled secondary antibody to target the primary antibody cocktail provides important practical advantages, because this way, the relative amount of fluorescent antibodies present on the cell surface is additive without also seeing an additive increase in the background signal as would happen with fluorescently conjugated primary antibodies (Figures 13 & **18**). We demonstrated the detection of CTCs that otherwise would be missed by targeting EpCAM alone with a dye-linked primary antibody. This immunolabeling strategy is an important improvement over our previously results as it addresses the heterogeneous nature of CTCs that makes them difficult to detect. The strategy also expands the limit of detection in a quantifiable manner, which is important for accurate CTC counts in patient samples.

(4) Six-fold increased recovery of CTCs from patient-derived samples. The relevance and success of the labeling schemes described above needed to be evaluated with clinical samples. To quantify the improvement, we compared the new labeling scheme to our former labeling strategy of PE-anti-EpCAM alone. **Figure 19** shows a bar graph with the number of CTCs isolated with each labeling scheme (cocktail and PE-anti-EpCAM) for each of the samples obtained from four breast cancer patients. Patient 1 was stage 3, ER+ PR+ HER2-; patient 2 was stage 4, ER+ PR- HER2+; patient 3 was stage 2, ER+ PR+ HER2-; and patient 4 was stage 2, ER+ PR+ HER2-. In this experiment, each sample was divided into two 2 mL portions. The samples were analyzed by eDAR sequentially using each of the two different labeling schemes. Samples were processed within two days after the blood was drawn. The data in Figure 19 demonstrate that a significant improvement, averaging 6×, was observed with the new labeling strategy. Figure 19B and 19C show representative cells that were isolated and further labeled against CD45, cytokeratin, and DAPI (post-capture) to confirm CTC identity.

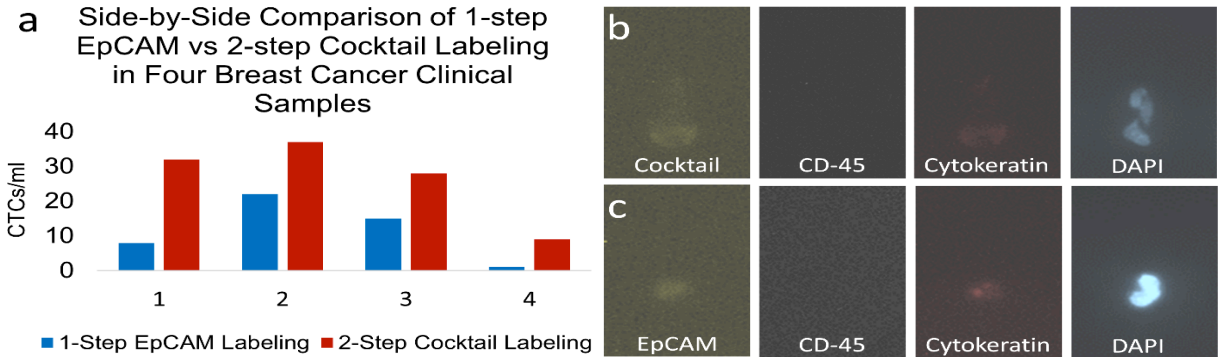


Figure 19. Cells from breast cancer clinical samples were sorted on an eDAR chip. Following sorting, the cells were fixed, permeabilized and labeled with CD45-alexa700, Pancytokeratin-alexa647 and DAPI to verify whether the cells were CTCs. A) Chart showing the number of CTCs recovered from each of three clinical breast samples, where 2 ml of blood was run with each 1-step EpCAM and 2-step cocktail. Patient 1 was stage 3, ER/PR+HER2-, patient 2 was stage 4, ER+PR-HER2+, and patient 3 was stage 2, ER/PR+HER2-. B) Fluorescence images of a cell captured by eDAR using the cocktail labeling scheme. C) A cell detected using the 1-step PE-anti-EpCAM labeling scheme.

These results demonstrate a significant improvement in CTC recovery with our new labeling scheme. They also validate the assertion that there are $EpCAM^{low}$ CTCs in the blood of breast cancer patients. We increased the capability to isolate and characterize these $EpCAM^{low}$ cells that represent a potentially more invasive population of CTCs. This enhanced sensitivity and selectivity can take advantage of our dual-filter sorting chip (described in last year's annual report) to isolate two distinct cell populations (high and low EpCAM expression) to separate on-chip filters. Finally, we are in the process of developing downstream analytical tools to extract genetic and phenotypic information from the CTCs (more details below in Task 2d). The information will be critical in creating a better diagnostic and prognostic tool for breast cancer as well as unraveling the mechanisms of metastasis and developing anti-metastatic drugs.

2d. Apply techniques developed for the prospective analysis of cancer cells (1a and 1c) from patient samples obtained in 2c with single-cell resolution (Months 36-60).

As described in previous reports, we have developed a manual method for placing single cells in individual wells and subsequently analyze each cell with digital nucleic acid and protein analysis. However, we found this method to be tedious when many single cells must be analyzed and has low-throughput. The current bottleneck in our ability to analyze single cells with high throughput arises from a lack of methods that can reliably and efficiently compartmentalize individual cells, each within a separate defined volume, such as placing each cell in a well of a 96-well plate.

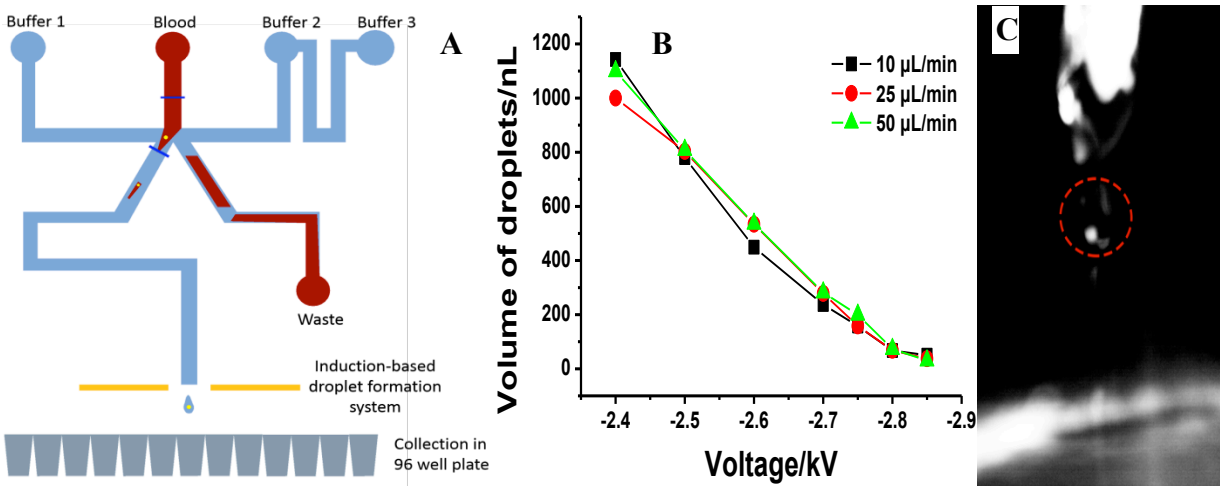


Figure 20. A) Schematic showing eDAR coupled to single-cell dispensing, where each individual cancer cell isolated by eDAR is then deposited into a well of a 96-well plate via single-cell dispensing using droplets formed by induction. B) A plot showing the volume of droplets as a function of the applied voltage and flow rate. C) An image of a droplet generated and in the process of being ejected under an applied voltage of 2.85 kV and a flow rate of 50 µL/min.

To address this issue, we worked to develop an automated single-cell dispensing technique, which can be coupled to eDAR (**Figure 20**) or as a standalone device. This technique relies on the use of induction to generate droplets in air, where each droplet deposited into the well would contain precisely a single cell. One initial difficulty with this induction method of droplet generation was the rather large size of the droplets produced, but we have since solved this problem and characterized the droplet diameter as a function of flow rate in the channel (Fig 20B). We are now able to form droplets as small as 30nL; Fig. 20C is an image showing the ejection of a droplet under an applied electric field.

At present, we are characterizing the performance of this method so we can determine the best parameter to use when we couple it to eDAR. **Figure 21** shows some of the preliminary results we collected, where we tried to understand how the size of the dispensed droplet containing the single cell is affected by the inner diameter of the capillary used, the composition and ionic strength of the aqueous solution, the flow rate in the capillary, and the applied voltage. We found the dispensed droplet volume is highly dependent on the composition of the aqueous solution as well as on the applied voltage, but not very sensitive to the inner diameter of the capillary or the flow rate in the capillary during dispensing. The flow rate, however, does have a large effect on the frequency of droplet dispensing and thus the duty cycle of the system. In addition to these operational parameters, we have also carried out preliminary experiments to characterize the viability of single cells contained in the droplet when ejected from the capillary nozzle. Here, we found the cells remain viable as long as the applied voltage is below a certain threshold, which is around 2.1kV under our current experimental conditions.

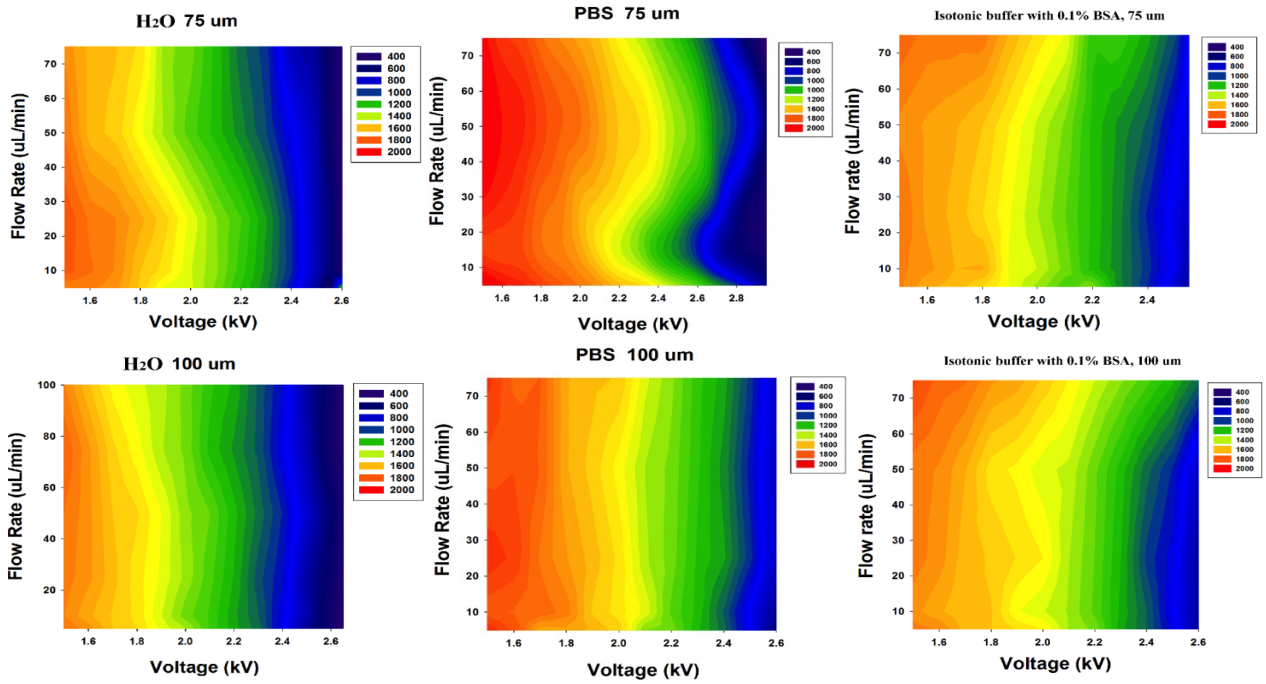


Figure 21. Color maps showing variations in the dispensed droplet size as a function of capillary inner diameter (75µm (top row) & 100µm (bottom row)), composition of aqueous solution (deionized water (left column)), phosphate buffer saline (PBS; middle column), & isoton with BSA (a blood diluent; right column)), flow rate in the capillary (y-axis), and applied voltage (x-axis). The colors show the droplet size as described in the accompanying color box, where the listed numbers have units of nLs.

Although the ejected droplets now have small dimensions and are suitable for use in single-cell dispensing, we found the droplets did not follow a straight path into a well placed underneath the nozzle after ejection (**Figure 22A**). This problem likely originates from the charges placed on the droplet during the ejection process. We have started to simulate the electric-field distribution around the device (Fig. 22B), which we believe will help us overcome this issue and which should allow us to focus the droplet containing the cell into the target well. We are also in the process of constructing and programming automated translation stages to work with this dispensing mechanism. When coupled to eDAR, this method would enable the high-throughput downstream genomic and proteomic analysis of each cell. As a standalone device, this method is useful for preparing single-cell samples for analysis, such as with single-cell SiMoA.

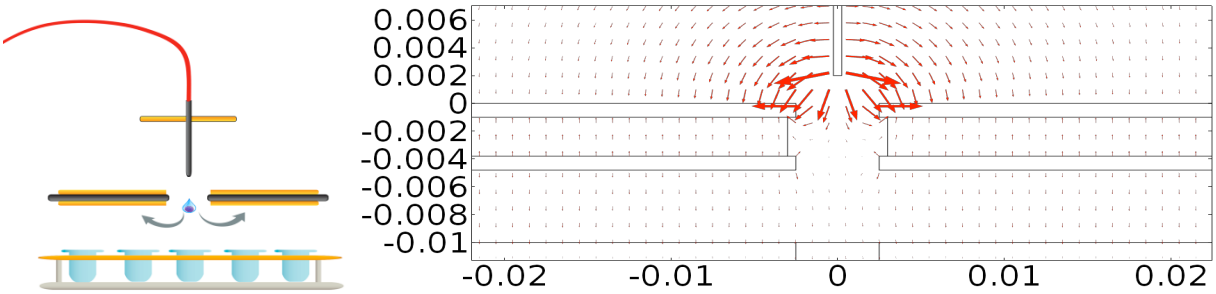


Figure 22. (Left Panel) Schematic showing the erratic path the droplet may take after ejection, likely caused by the charges placed on the droplet during ejection. (Right Panel) Simulation showing the electric-field distribution of the current device shown in the left panel, where the capillary was at 2kV, the top plate (0 on y-axis) was at ground, the bottom plate (-0.004 on y-axis) was at 250V, and the well plate was at 500V.

Charlotte Kuperwasser, *Tufts University School of Medicine, Boston, MA*

Task 1 Early Detection (Months 1 to 36)
1a. Utilize the 'Human-In-Mouse' breast tumor model to determine if one can identify biomarkers of early tumor progression and tumor growth. Collect blood and breast tissues at various stages of progression: Normal, hyperplasia, DCIS, invasive cancer (Months 1-18).
HIM tumors will be created from human breast epithelial cells collected from discarded tissues of women who have undergone reduction mammoplasty surgeries. For more information/details on model: *Nature Protocols*, **2006**, 1, 595-599.

As described in the quarterly reports for this year, Task 1 has been completed.
We have generated nearly 20 tumors and have provided the Walt lab with all the samples they have requested for their studies.

In addition, while the animal studies and generation of HIM tumors have been valuable for the studies by the Walt and Sonenshein laboratories, at this stage of the project, testing the identified and validated markers from HIM samples will be focused in human samples. Thus, this year we plan to generate HIM tumors and even some PDX tumors for the Walt lab's single cell studies.

Task 2. Response to chemotherapy and hormone therapy (Months 24-48).

2a. Inject traditional breast cancer cell-line based xenografts into mice using hormone receptor positive lines and hormone receptor negative and Her2+ lines (MCF7, T47D, HCC1418, SUM225, BT20, SUM149, SUM159)(Months 24-36).
Human breast cancer cell lines are all commercially available and previously characterized. For more information/details on these cell line-based xenografts see: *Breast Cancer Research*, **2010**;12(5):1-17.

Cell-line based xenografts have been initiated and terminal blood collections have been made with MCF7, T47D and SUM1315 lines (N=10, 10, 12, respectively). Chemotherapy and hormone therapy treatments and studies have not yet been initiated.

2b. Allow tumors to form and reach 5mm and assess serum levels of EMT/CSC markers. (Months 24-36)

2c. Treat animals with chemotherapy (Paclitaxel/Taxol) or anti-estrogen (Tamoxifen) and measure serum levels of EMT/CSC markers and mtDNAs during tumor regression and after cessation of therapy.

2d. Continue to monitor tumor growth. (Months 36-48)

2e. Determine whether the levels of EMT/CSC markers correlate to response to therapy or to tumor recurrence (Months 36-48). Predicted Total mice =(30x2replicates)=60

These studies have not yet been initiated per SOW.

Task 3. Recurrence prediction (Months 36-60)

3a. Create tumors using the 'Human-In-Mouse' breast tumor model. Allow tumors to form and reach 5-8mm in diameter and assess serum levels of oncogenes used to create tumors (Months 36-50).

3b. Surgically resect the tumors and measure blood serum levels of EMT/CSC markers as well as mitochondrial DNAs during tumor regression and after cessation of therapy after surgery (Months 36-50).

3c. Monitor mice and measure the levels of EMT/CSC markers weekly to determine if they increase or correlate with regrowth of tumor (Months 48-60). Predicted Total mice=(80x2replicates)=160;Predicted total number of reduction mammoplasty tissues= 10

These studies have not yet been initiated per SOW.

Gail E. Sonenshein, Ph.D., *Tufts University School of Medicine 150 Harrison Avenue Boston, MA 02111*

Task 1. Perform mitochondrial DNA mutational analysis. (Months 1 to 16).

1a. Isolate mitochondria from cultured mouse and human mammary breast epithelial or cancer cells that exist in our laboratory and extract DNA and subject it to massively parallel sequencing and bioinformatics to identify mutations. (Months 1-8).

As mentioned in our previous annual report, we have developed the technologies and used these to characterize the mitochondrial DNA sequences in human and mouse cell lines and tumor tissue. Thus this task is essentially completed. However, as new inflammatory breast cancer (IBC) cell lines are established or as we acquire additional lines, we are analyzing their mtDNA sequences to increase our database.

1b. Isolate mitochondria from normal tissue and from tumors, extract DNA and subject it to massively parallel sequencing (Months 8-12).

We have developed methods to extract DNA from frozen and formalin fixed paraffin embedded (FFPE) tumor tissue. We have been able to successfully amplify DNA obtained from these tissues and are now ready to analyze tumor tissue obtained from the HIM breast tumor model. In particular, we are working on SUM-1315 tumor tissue obtained from the Kuperwasser lab. Our analysis has led us to identify key mutations observed repeatedly in both breast cancer cell lines and tumor tissue (**Table 7**). These key mutations may be useful in detecting breast tumors in patients in the future.

1c. Compare mtDNA from normal mouse mammary tissue and from tumor tissue using bioinformatics to identify mutations. (Months 10-16).

As mentioned in the previous annual report, we were previously having some difficulty in designing primers for mouse DNA as those reported in the literature produced large PCR products not amenable for sequencing. We have now identified and confirmed our new mouse specific primers and have started analyzing mouse DNA. We will next analyze DNA from transgenic mice developed in the Kuperwasser lab (Task 2a).

Mutation	Gene	Mutation	Gene
73 A>G	D-LOOP	7028 C>T	CO1
152 T>C	D-LOOP	7521 G>A	CO1
185 G>A	D-LOOP	8764 G>A	ATPASE6
189 A>G	D-LOOP	8701 A>G	ATPASE6
195 T>C	D-LOOP	8860 A>G	ATPASE6
198 C>T	D-LOOP	9540 T>C	COIII
228 G>A	D-LOOP	10873 T>C	ND4
263 A>G	D-LOOP	10398 A>G	ND4
295 C>T	D-LOOP	11251 A>G	ND4
311 INSERT C	D-LOOP	11719 G>A	ND4
489 T>C	D-LOOP	12612 A>G	ND5
523 A-DEL	D-LOOP	12705 C>T	ND5
750 A>G	12s	13650 C>T	ND5
769 G>A	12S	13263 G>A	D-LOOP
1438 A>G	12S	15326 A>G	D-LOOP
2706 A>G	16S	16278 C>T	D-LOOP
3010 G>A	16S	16294 C>T	D-LOOP
4216 T>C	ND1	16311 T>C	D-LOOP
4769 A>G	ND2	16319 G>A	D-LOOP
3594 C>T	ND1	16519 T>C	D-LOOP

Table 7. Common mutations observed in breast cancer cell lines and tissues. All mutations highlighted in **red** are high frequency mutations not specific to breast cancer and in **black** are low frequency mutations more specific to breast cancer.

Task 2. Characterize mtDNA mutations resulting from oncogene expression in breast tumors (Months 16-24).

2a. Compare mtDNA from mammary tissue vs tumors or derived cell lines of transgenic mice driven by oncogenes implicated in breast cancer. Samples obtained from Kuperwasser (months 16-24)

We obtained unmatched serum and tissue samples from HIM tumors with SV40/KRas mutation from the Kuperwasser lab. We have identified 32 mtDNA mutations induced by the K-*Ras* oncogene in these tumors (**Table 8**). We also looked for these mutations in the matched serum samples obtained from mice. Since the serum samples collected at each time point were very small, we pooled samples. Unfortunately, even after pooling the sample amount was not enough to isolate adequate amounts of mtDNA. We worked with Dr. Walt's group to try to setup a single molecule array (SiMoA) platform for amplifying small amounts of mtDNA mutations. However, the efforts were not successful in developing this method likely due to the small differences in sequence. So in the future, we will use blood rather than serum samples, which permits amplification.

Kras tumor

Gene	nucleotide position	mutation
D-Loop	146	T>C
D-Loop	195	T>C
D-Loop	263	A>G
D-Loop	310	insert C
D-Loop	317	insert C
12sRNA	709	G>A
12sRNA	750	A>G
12sRNA	775	C>A
12sRNA	794	insert C
12sRNA	1438	A>G
16sRNA	1793	G>A
16sRNA	3248	insert G
ND1	3437	g>c
ND1	3692	g>c
ND1	4083	g>A
ND2	4700	inserta
ND2	4769	A>G
ND2	5095	T>C
tRna-cys	5773	G>T
COX1	6218	insert t
COX1	6572	insert G
COX2	8107	A>G
ATPASE6	8764	g>a
ATPASE6	8860	A>G
COX3	9812	c>t
ND5	13102	a>c
ND5	14122	c>t
CYB	15016	insert t
CYB	15326	A>G
D-Loop	16068	T>C
D-Loop	16288	T>C
D-Loop	16362	T>C

Table 8. DNA gel demonstrating clean amplification of mtDNA in the SV40/K-*Ras* tumors obtained from the Kuperwasser lab

We were also interested in determining whether BRCA1 carriers have particular mtDNA mutations. Breast tumors from patients with the BRCA1 mutation are very difficult to detect by mammography and these patients are more susceptible to irradiation. Thus, we have analyzed DNA from mice carrying HIM tumors with BRCA1 mutations. We identified 31 mutations in the BRCA1 tumor sample (**Table 9**). Interestingly, we found a tRNA mutation in this tumor tissue, which is very rare and unique. Of note, this mutation, A12308G, has been reported to correlate with breast cancer risk. Since we observe characteristic mutations in the tumor DNA, we will assess blood samples to determine whether a blood test for mtDNA mutations can be developed. We will also analyze more tumor samples to determine whether this tRNA mutation is characteristic of BRCA1 disease and to identify other characteristic mtDNA mutations.

gene	np	mutation
D-loop	16327	C>T
D-Loop	16519	T>C
D-Loop	73	A>G
D-Loop	146	T>C
D-Loop	263	A>G
D-Loop	285	T>C
D-Loop	315	ins c
12sRNA	1438	A>G
16sRNA	2387	T>C
16sRNA	2706	A>G
ND1	3425	A>G
ND2	4769	A>G
ND2	5460	G>A
COX1	6365	T>C
COX1	7028	C>T
ATPASE8	8395	C>T
ATPASE6	8860	A>G
	10469	insert G
ND4	10886	T>C
ND4	11467	A>G
ND4	11566	A>G
ND4	11719	G>A
MT-TL2	12308	A>G
ND5	12372	G>A
ND5	12879	T>C
ND5	13104	A>G
CYB	14766	C>T
CYB	15110	G>A
CYB	15148	G>A
CYB	15172	G>A
CYB	15326	A>G

Table 9. Mutations observed in BRCA1 positive HIM tumor tissues. Mutations highlighted in yellow are commonly observed in breast cancer cell lines and patient samples. Mutation highlighted in blue is a tRNA mutation reported to be associated with breast cancer risk.

2b. Characterize mtDNA alterations in tumor lines that will be utilized by the Kuperwasser laboratory in their 'HIM' analysis. DNA samples will be analyzed as in Task 1 above (Months 18-24).

We have obtained tumor tissue from the HIM tumors derived from the SUM1315 breast cancer cell line. DNA from these tissues has been amplified and was found to have no new mutations when compared to the SUM1315 cell line. These data along with other data obtained from our lab suggest that under the conditions being used, the tumor microenvironment is not inducing additional mtDNA mutations during tumor progression.

Task 3. Recurrence prediction: Determine whether mtDNA mutations can be used to test for breast tumor regression in mice (in collaboration with Kuperwasser) (Months 18-60).

3a. mtDNA from tumor samples provided by the Kuperwasser laboratory using the HIM breast tumor model. Tumors will be initially removed after they have been allowed to reach 5-8 mm in diameter and then at the end of the experiment (Months 18-30)

As mentioned previously, we are using technology developed in our lab to assess whether mtDNA in the blood can be used to evaluate tumor regrowth/recurrence. Briefly, NOD/SCID mice were injected with 2.5 x 106 MDA-MB-231 cells in the mammary fat pad. Mice were bled a day prior to injection and on days 7 and 16 after cancer cell inoculation. The tumors were resected when they reached an average volume of 0.2 cm^3. Mouse 3 died before resection. As shown in **Table 10** below, tumor resection was successfully performed on 11 mice, which were then followed for a total of 32 days. Out of the 11 mice, 6 mice had primary recurrence over the 5-week period. All of the mice developed macro- or micro-metastases to the lung, including the ones that had no primary recurrence. Four of the mice developed brain metastases (see Table 10). We are currently using whole blood and tissue samples obtained from our resection study mentioned in the annual report of our work and have addressed the questions below.

Mouse No#	# 1	# 2	#3 Died	# 4	# 5	# 6	#7	#8	#9	#10	#11	# 12
Primary recurrence	+ 3/20/14 17 days after resection	+ 4/03/14 33 days after resection		+ 3/7/14 4 days after resection	+ 3/7/14 4 days after resection	None	None	+ 3/14/14 11 days after resection	+ 3/7/14 4 days after resection	None	None	None
Brain metastasis	+m	+m		ND	+m	–	+	–	ND	–	–	–
Lung metastasis	+m	+		ND	+	+m	+	+	ND	+	+	+

Table 10. This table depicts a summary of the resection experiment, including information of the specific mice that showed primary recurrence, brain metastases or lung metastases. +m: multiple metastases, +: single metastases, –: no metastases, ND: not determined. Most of the metastases observed in the brain were smaller micro metastases and those observed in the lung were macro metastases.

Can mtDNA be observed in mice with early recurrence?

Yes. We previously identified mtDNA in blood of mice 4 and 9 that showed recurrence 4 days after resection. We have now expanded our study to include mice 5 and 8, which showed recurrence at a slightly later date (**Table 11**). We have now been able to identify mtDNA mutations in the blood of both of these mice.

Can we observe mtDNA mutations in mice with no primary recurrence but lung and brain metastases?

Our data show that we can detect mtDNA mutations in blood of mice that show recurrence (Table 11). We have now analyzed the blood of mice that did not show recurrence like mice 6, 7, and 12 to determine if we can predict tumor recurrence before it happens. We were able to detect mtDNA mutations in mice with no primary recurrence as early as 10 days after resection. To confirm that the presence of mtDNA mutation was from the metastasis and not from the residual cells left from the tumor, we analyzed blood obtained from these mice 17 days after resection (Table 11). Our results show that we can identify mtDNA mutations in the blood of all mice tested; thus, testing for specific mtDNA mutations, which are markers of disease, holds potential for early detection of breast cancer recurrence in patients.

Can we establish a panel of mtDNA mutations for early detection and recurrence?

This question is currently under investigation. We will establish a panel of mtDNA mutations based on our recurrence study and test whether they can accurately identify recurrence and metastasis in a new study.

MDA_MB-231			Observed in blood of mice 4 and 9	Observed in blood of mice 5, 8 and 7 on 3/13/14 10 days after resection (AFR)	Observed in blood of mice 7, 6, 12 on 3/13/14 10 days AFR	Observed in blood of mice 7, 12, 6 3/20/14 17 days AFR
gene	np	mutation				
D-Loop	16093	T>C	yes			
D-Loop	16189	T>C	yes			
D-Loop	16265	A>G				
D-Loop	16278	C>T				
D-Loop	16519	T>C				
D-Loop	73	A>G	yes	yes		
D-Loop	153	A>G	yes	yes		
D-Loop	195	T>C	yes	yes		
D-Loop	263	A>G	yes	yes		
D-Loop	309-311	ins c	yes			
D-Loop	523-524	INS C				
12sRNA	709	G>A				
12sRNA	750	A>G	yes			
12sRNA	1438	A>G				
16sRNA	1719	G>A				
16sRNA	2706	A>G				
ND2	4769	A>G	yes			
COX1	6221	T>C	yes	yes	yes (7,6)	
COX1	6371	C>T	yes	yes	yes (7,6)	
COX1	7028	C>T				
ATPASE8	8506	T>C	yes			
ATPASE6	8860	A>G	yes			
ND4	11719	G>A	yes			
ND4	12084	C>T	yes			
ND5	12705	C>T	yes			yes (12, 7,6)
ND5	13966	A>G				
ND6	14470	T>C	yes			
CYB	14766	C>T	yes	yes		
CYB	15310	T>C	yes	yes	yes (6)	yes (12)
CYB	15326	A>G	yes	yes	yes (6)	yes (12)
D-loop	16223	C>T				

Table 11. Mutations observed in blood samples of mice with recurrence collected 10 or 17 days after resection. All mutations highlighted in red are observed in blood samples of mice 4 and 9, ten days after resection. Some of these mutations were also observed in 5, 7, 8, 6 and 12. This was the earliest time point at which we collected blood after resection.

3b. DNA will be isolated from serum collected before resection and weekly after resection that will be provided by the Kuperwasser laboratory (Months 18-30).

In our last annual report, we demonstrated that we can detect mtDNA in the blood of mice at a stage when the tumor is barely palpable, when circulating tumor cells (CTCs) are detectable. Together, these results suggest that mtDNA mutations can be used as markers for detection of early stage breast cancer recurrence in patients. We are currently testing blood samples obtained from the Kuperwasser lab from SUM1315 derived HIM tumors and are also looking for an increase in the frequency of mutations with tumor progression. However, the average serum volume size is less than 20 µl per time point and is insufficient for DNA extraction using our protocol. We have tried to pool some samples from the same time point, but the DNA obtained has still been insufficient for PCR. We tried using Blood Phusion direct PCR kit (Thermo Scientific), which can amplify small amounts of DNA directly from blood without the need to isolate the DNA. This kit was unable to amplify DNA from whole blood. As discussed above in Task 3a, we are able to successfully test whole blood samples obtained after resection from mice injected with MDA-MB-231 cells, but have found very little to no mtDNA in serum samples. Thus, we propose in the future to focus on whole blood samples, which contain more mtDNA.

In summary, in the last year we have identified common mutations observed in breast cancer that can be used as potential markers for early detection and recurrence. We identified novel mtDNA mutations induced by KRAS and BRCA1 mutant proteins leading to breast cancer. Specifically, we have identified a unique tRNA mutation in a BRCA1 mutation carrying tumor sample, which has been linked to increased breast cancer risk in patients, thus justifying further study into the role and importance of tRNA mutations in BRCA1 carriers as markers for increased risk and early detection. We have also successfully completed our mouse experiment and have answered important questions about the detection of mtDNA mutations during tumor progression and recurrence. We have shown that we can detect mtDNA mutation as early as 10 days after resection when there is no primary recurrence. Most importantly, we can detect mtDNA mutation in the blood of mice that have no primary recurrence but have lung and brain metastases.

Rachel Buchsbaum, Tufts Medical Center, 75 Kneeland St., Boston, MA 02111

The Buchsbaum research effort includes both laboratory and clinical research aims. The purpose of the laboratory effort is to determine the effect of the tumor microenvironment on serum markers of metastasis in the HIM model. The purpose of the clinical research effort is to validate the serum markers of metastasis derived from the Walt and Chiu pre-clinical experiments using human serum samples.

Task 1. Determine effect of tumor microenvironment on serum markers of metastasis in the HIM model (Months 1-36).

1a. Develop mammary fibroblast lines with range of Tiam1 (both down- and up-regulated). (Months 1-6)

Completed as detailed in prior reports.

1b. Utilize HIM breast tumor model in collaboration with Dr. Kuperwasser.
In earlier Annual Progress Reports, we summarized our analysis of the effects of Tiam1-manipulated fibroblasts in the 3D in vitro breast cancer microenvironment, as well as two mouse models of human breast cancer metastasis. We also showed initial results using a chemical inhibitor (here called inhibitor 1) to the Tiam1-deficiency pathway and its effect in blocking metastasis in the second mouse model. In this model, the mice receive xenografts with breast cancer cells isolated from 3D co-cultures after exposure to engineered mammary fibroblasts. In the original protocol, we used inhibitor 1 to treat 3D tissue cultures, prior to xenograft implantation in the mice.

During the prior year, we have initiated a modification to the second mouse model of human breast cancer metastasis. In the new modified model, mice are treated directly with the inhibitor, rather than treating the co-cultures with the inhibitor. To do this, we worked with our supplying collaborator to derive a suitable supply of the inhibitor. We applied for and received approval from our institutional DLAM for approval to amend our current animal protocol to permit this treatment. At the end of the prior year, and continuing in the current reporting year, we initiated the 3D tissue cultures necessary for the implantation experiments, and treated sixteen additional mice on this modified protocol. All mice developed primary tumors at the site of implantation, and received direct dosing with inhibitor 1, consisting of daily i.p. injections beginning approximately two weeks after implantation (once staples were removed). Mice tolerated the treatment for only 2-3 weeks before developing weight loss and ruffled fur, requiring sacrifice one-two weeks earlier than usual. Terminal serum specimens were obtained on all mice. Pathology studies were obtained on tumors and lungs, confirming primary tumors and showing no clear lung metastases. This is likely due to the early sacrifice required by developing toxicity to inhibitor 1 treatment.

Thus, we have focused on the development of more potent, less toxic inhibitors, together with our collaborator. From an earlier screen of six additional related compounds testing potential inhibitory ability against the Tiam1-deficiency pathway, we have now identified two potential novel inhibitors. During the earlier reporting period, these screening tests were complicated by varying issues affecting the reproducibility of results. These have now been largely resolved. During the current reporting period we have focused on one of the two novel inhibitors (here termed inhibitor 2). We have now reproducibly demonstrated several-fold increased potency of inhibitor 2 in both the screening studies as well as more complex first and second-phase functional assays of migration, invasion, and cancer stem cell populations. We have now also gone on to test 15 derivative compounds synthesized by our collaborator based on the findings with inhibitors 1 and 2. We have identified active inhibition of the Tiam1-deficiency pathway in

3 of these derivatives, which include derivatives of both inhibitors 1 and 2. We are currently conducting expanded dose titration experiments in vitro to establish optimal conditions for testing this new compound in functional studies based on the 3D culture model. The goal is to establish an optimized inhibitor compound for testing in the mouse xenograft models. As before, serum samples will be obtained and banked during all mouse experiments for potential marker assays.

Finally, during this reporting period we also initiated experiments testing the effect of our inhibitor 1 in an expanded phenotypic range of breast cancer cell lines. Our original tests were performed using triple-negative breast cancer lines. We have found that two hormone-receptor positive cell lines formed spheroids in 3D co-culture with our engineered fibroblasts, and that exposure to Tiam1-deficiency fibroblasts enhanced cancer cell invasion into matrix. Experimental results with these two new lines suggest that incorporation of inhibitor 1 blocks the effect of Tiam1-deficient fibroblasts on cancer cell invasion and mammosphere formation. This suggests that the Tiam1-deficiency pathway in breast cancer-associated fibroblasts can affect multiple different sub-types of breast cancers, thus increasing the potential clinical impact of these studies.

Task 2. Perform retrospective clinical trial to validate candidate markers from Walt single molecule studies using banked repository samples. (Months 1-60)

During the previous reporting periods, we worked with statisticians in the Tufts Research Design Center within the Tufts Clinical and Translational Sciences Institute to define needed numbers of human serum samples for testing of the initial biomarker panel, based on prevalence estimates derived from reports of circulating tumor cells.

We also identified potential samples within the NIH PLCO (Prostate, Lung, Colorectal and Ovarian) Cancer screening trial. In particular, this trial includes samples from women prior to development of breast cancer, which are otherwise very difficult to find in suitable number and sample volume. A Concept Application was submitted to the PLCO as the initial step in the process for access to the samples, based on recent data from the Walt and Kuperwasser labs derived from proof-of-concept experiments demonstrating the ability to detect circulating PSA in mice implanted with human prostate cancer cells prior to the development of measurable tumors.

During this reporting period, we received notification that our application to the EEMS-PLCO trial was denied for the winter cycle. The primary reason was the volume of serum sample required. Secondary concerns were the justification of the request for post-diagnosis serum samples and a request to see more details of the preliminary data. At the current time, the EEMS-PLCO is not accepting new applications due to internal restructuring. We anticipate resubmission of a revised application once the application process is again open. The Walt lab is working on reducing the sample volumes required.

During this reporting period, we also continued our search for other suitable biorepositories for provision of serum samples. This search has included extensive review of many of the biorepository resources listed through the NCI Specimen Resource Locator (https://specimens.cancer.gov/) as well as other biorepositories. To date, we have identified two other potential sources of specimens suitable for this project: the Baltimore Longitudinal Study of Aging (http://www.blsa.nih.gov/) and the Gundersen Foundation BioBank (http://www.gundersenhealth.org/biobank). The defined criteria for sample suitability include the ability to obtain accompanying curated clinical information regarding tumor status and date of sample acquisition with respect to treatment.

Multiple phone calls to the BLSA requesting further information about access to samples have not been returned. However, our initial application to the Gundersen Foundation has identified over 130 suitable samples for this project in a search of their biorepository. Access to these samples will require significant additional cost, so these samples have not yet been obtained.

We have initiated the request for verification to our institutional IRB regarding our judgment that working with de-identified serum samples from public biobanks such as the EEMS-PLCO and the Gundersen Foundation BioBank does not constitute Human Subjects Research. Our IRB has confirmed our judgment that working with de-identified serum samples from public biobanks such as these does not constitute Human Subjects Research. Our IRB has requested that we obtain confirmation from the biobanks that samples will be de-identified, prior to accessing any such samples.

Task 3. Perform prospective clinical trial for predictive and prognostic markers in women with newly diagnosed breast cancer.

Progress on developing a bank of human serum samples for validation of candidate markers is continuing. The goal is to establish a fully curated bank of serum samples from all women with newly diagnosed breast cancer, women who are being followed after completing treatment for localized breast cancer, and women with active metastatic breast cancer being treated at Tufts Medical Center. Particular focus has been on banking serum from women with newly diagnosed breast cancer, prior to definitive surgery, as these samples are not generally available from outside serum collections. Recruitment of consenting subjects and serum samples to our Tissue Repository is currently ongoing. During this period, we have also successfully recruited other clinicians to this effort, most importantly breast cancer surgeons and breast center personnel. The resulting significant increase in consenting subjects and requests for sample collections exposed a number of logistical issues in the process of handling specimens in the main laboratory prior to delivery to the Tissue Repository. During the current reporting period, substantial progress was made in identifying and addressing these process issues. At this time, we have successfully collected over 170 samples. The success rate of obtaining and banking ordered samples from consenting subjects has improved to 96% during this reporting year (135/140 samples) from 64% previously (41/64 samples).

KEY RESEARCH ACCOMPLISHMENTS

- Developed SiMoA assays for several breast cancer protein biomarkers.
- Established a method to detect and distinguish miRNAs with high sensitivity and specificity.
- Performed preliminary studies to quantify candidate biomarker levels in healthy and diseased human serum and evaluate the usefulness of each marker.
- Characterized multiple breast cancer cell lines with established SiMoA assays.
- Measured PSA content of two different populations of single LNCaP cells and observed differences based on protein expression.
- Isolated CTCs from early stage breast cancer patients using eDAR.
- Characterized the performance of eDAR.
- Developed ultra-bright probes for the isolation of CTCs with low biomarker expression.
- Improved the ability of eDAR to capture CTCs with variable patterns of biomarker expression.
- Demonstrated a six-fold increase in recovery rate of CTCs from patient-derived samples.
- Explored high-throughput methods for single-cell dispensing for downstream analysis of CTCs using digital nucleic acid and protein analysis.
- Experimentally confirmed that no mutations were induced in vivo when tumor cells are injected in NOD/SCID mice or in HIM mouse models.
- Detected mtDNA mutations as early as 7 days after tumor cell injection when there is no palpable tumor in mice.
- Detected mtDNA mutations after surgery in mice with no primary recurrence but with lung and brain metastases.
- Detected mtDNA mutations in mice early on before the primary tumor recurs.

- Identified new compound with increased potency against metastasis and decreased toxicity based on in vitro screening and functional assays.
- Identified new derivative compounds with potential activity against metastasis based on in vitro screening.
- Established expanded range of breast cancer sub-types with sensitivity to our initial inhibitor.
- Identified available human serum banks for biomarker testing, including the Gundersen Foundation BioBank and NIH-PLCO screening trial. Initial applications were submitted to both. Now awaiting resolution of funding issues (Gundersen) and re-opening of application process (PLCO) in order to proceed.
- Established effective process for serum banking from women with breast cancer at Tufts Medical Center; active monitoring ongoing.

REPORTABLE OUTCOMES

- Schubert, S. M.; Arendt, L. M.; Zhou, W.; Baig, S.; Walter, S. R.; Buchsbaum, R. J.; Kuperwasser, C.; Walt, D. R., "Ultra-sensitive protein detection via Single Molecule Arrays towards early stage cancer monitoring", *Scientific Reports* **2015,** *5*, 11034.
- Schubert, S. M.; Walter, S. R.; Manesse, M.; Walt, D. R. "Protein Counting in Single Cancer Cells", Manuscript Submitted.
- Ghatak, P.; Hartman, M. R.; Walt, D. R. "Detection of microRNAs as Potential Biomarkers of Early Stages in Breast Cancer". Poster presentation. Extracellular Biomarkers Summit, 2015 March 16, Cambridge, MA.
- Baig, S.; Schubert, S. M.; Walter, S. R.; Usmani, K.; Walt, D.R. "Ultrasensitive Assays for Early Breast Cancer Detection". Poster presentation. The Pittsburgh Conference on Analytical Chemistry and Applied Spectroscopy, 2015 March 8-12, New Orleans, LA.
- Baig, S.; Schubert, S. M.; Walter, S. R.; Walt, D. R. "Single Molecule Assay Development for Breast Cancer Detection". Poster presentation. 250th American Chemical Society National Meeting and Exposition, 2015 August 16-20, Boston, MA.
- Schubert, S. M.; Baig, S.; Walter, S. R.; Palacios, M.; Walt, D. R. "Development of Serum-Based Single Molecule Assays for the Early Detection of Cancer." Oral Presentation. 250th American Chemical Society National Meeting and Exposition, 2015 August 16-20, Boston, MA.
- Walter, S. R.; Schubert, S. M.; Manesse, M.; Walt. D. R. "Quantifying Protein Expression in Single Cells." Oral Presentation. 250th American Chemical Society National Meeting and Exposition, 2015 August 16-20, Boston, MA.
- A.M. Thompson, A. Gansen, A.L. Paguirigan, J.E. Kreutz, J.P. Radich, D.T. Chiu, "The self-digitization microfluidic chip for the absolute quantification of mRNA in single cells" *Anal. Chem.,* **2014,** 86, 12308-12314.
- R.K. Anand, E.S. Johnson, D.T. Chiu, "Negative dielectrophoretic capture and repulsion of single cells at a bipolar electrode: the impact of faradaic ion en-richment and depletion" *J. Amer. Chem. Soc.*, **2015**, 137, 776-783.
- M. Zhao, B. Wei, W.C. Nelson, D.T. Chiu, "Simultaneous and selective isolation of multiple subpopulations of rare cells from peripheral blood using ensemble-decision aliquot ranking" *Lab Chip*, **2015**, 15, 3391-3396.
- E.S. Johnson, R.K. Anand, D.T. Chiu (2015) "Improving detection of circulating tumor cells with low numbers of the targeted surface antigen by ensemble decision aliquot ranking (eDAR)" *Anal. Chem.* 87, 9389-9395.
- M. Zhao, P.G. Schiro, D.T. Chiu (2016) "Ensemble-decision aliquot ranking (eDAR) for CTC isolation and analysis" *Circulating Tumor Cells: Isolation and Analysis* (in press).

- S. G. Das, M. Romagnoli, N. D. Mineva, S. Barillé-Nion, P. Jézéquel, M. Capone, G. E. Sonenshein, "Identification of miR-720, as a downstream target of ADAM8 promoting migratory and invasive phenotype in triple negative breast cancer", submitted.
- Sonenshein, G.E. "ADAM8, a Driver of Metastasis, is a Novel Target for Treatment of Triple-Negative Breast Cancer". Oral presentation. Northeastern Breast Cancer Research Conference, 8 May 2015, Lebanon, NH.

CONCLUSION

In conclusion, great progress was made in the development of SiMoA assays for the targeted breast cancer biomarkers. A total of 13 SiMoA assays have now been successfully completed, most with LODs significantly lower than their commercial ELISA counterpart. Although some of these assays can still be optimized to gain more sensitivity, more focus will be placed on utilizing these assays for determining biomarker concentrations in breast cancer patient samples. We have also begun preliminary screening of multiple biomarkers in breast cancer patient and healthy control samples. Although additional samples must be screened to gain greater statistical information and further data processing is required, our preliminary results seem promising. Additional markers, such as mtDNA will be examined to improve the sensitivity and specificity of the SiMoA assays.

In addition, we have established a mouse model to highlight the use of SiMoA for the early detection of cancer. This proof-of-concept study using PSA demonstrated that biomarkers can be measured at ultrasensitive levels within serum using SiMoA prior to tumor formation. These results are vital to the implementation of SiMoA technology as an early detection and monitoring platform for breast cancer.

We have successfully achieved full quantification of protein molecules within single cancer cells using SiMoA. Highly cultured cells were found to exhibit different protein expression levels than low-passage cells. This technique will be applied to quantify protein copy numbers in single breast cancer cells using several of the established SiMoA assays described above. We are now attempting to integrate eDAR with SiMoA to increase the throughput of the system.

REFERENCES

(1) Wilson, D. H.; Hanlon, D. W.; Provuncher, G. K.; Chang, L.; Song, L.; Patel, P. P.; Ferrell, E. P.; Lepor, H.; Partin, A. W.; Chan, D. W.; Sokoll, L. J.; Cheli, C. D.; Thiel, R. P.; Fournier, D. R.; Duffy, D. C. *Clinical Chemistry* **2011**, *57*, 1712.